J2UL-1207-08Z0(00)

FUJITSU Server PRIMEQUEST 1000 シリーズ

Linux ユーザーズマニュアル

Red Hat Enterprise Linux 5 編

FUJITSU

本書をお読みになる前に

本書は、以下の対象機種で Red Hat Enterprise Linux 5 **をお使いになる方を対象に、設計／導入／運用／保守に関する情報や参考となる考え方を提供しています。導入から運用／保守をスムーズに実施するため、各段階において本書をご活用ください。**
なお、FUJITSU Server PRIMEQUEST 1000 **シリーズで仮想マシン機能を利用される場合は、『**PRIMEQUEST(1000 **シリーズ**) RHEL5-Xen **仮想マシン機能ユーザーズマニュアル』を参照してください。**

● **対象機種**

FUJITSU Server PRIMEQUEST 1000 シリーズ

本書の位置づけ

本書は Red Hat 社から公開されている Red Hat Enterprise Linux 5 向けマニュアルに対する補足情報を提供します。
Red Hat Enterprise Linux 5 の機能や使い方の説明については、はじめに Red Hat 社が公開するマニュアルを参照してください。Red Hat 社が公開しているマニュアルについては、本書の第 1 章で参照方法を説明しています。

高度な安全性が要求される用途への使用について

本製品は、一般事務用、パーソナル用、家庭用、通常の産業などの一般的用途を想定して開発・設計・製造されているものであり、原子力施設における核反応制御、航空機自動飛行制御、航空交通管制、大量輸送システムにおける運行制御、生命維持のための医療用機器、兵器システムにおけるミサイル発射制御など、極めて高度な安全性が要求され、仮に当該安全性が確保されない場合、直接生命・身体に対する重大な危険性を伴う用途（以下「ハイセイフティ用途」という）に使用されるよう開発・設計・製造されたものではありません。
お客様は本製品を必要な安全性を確保する措置を施すことなくハイセイフティ用途に使用しないでください。また、お客様がハイセイフティ用途に本製品を使用したことにより発生する、お客様または第三者からのいかなる請求または損害賠償に対しても富士通株式会社およびその関連会社は一切責任を負いかねます。

本書の構成

本書の構成は以下のとおりです。

章	内容
本書をお読みになる前に	本書で使用している記号や表記の説明、参照先ドキュメントなどの情報を記載しています。
第 1 章　概要：Linux のノウハウ	この章では、Linux をお使いになるうえで知っておいていただきたい情報と、Linux ご使用の流れについて説明します。
第 2 章　システム設計：ハードディスク	この章では、ハードディスク関連の設計情報を説明します。
第 3 章　システム設計：ネットワーク	この章では、ネットワークの設計に関する設定方法や注意事項について説明します。
第 4 章　システム設計：デバイス名ずれ対策設計	この章では、デバイス名ずれの要因およびデバイス名ずれに対応するために必要な設計について説明します。
第 5 章　システム設計：ログファイル運用	この章では、ログファイル運用について説明します。
第 6 章　システム設計：時刻補正の導入（NTP）	この章では、Linux システムにおける時刻の仕組みや時刻補正の必要性、時刻補正に重要な役割を果たす NTP（Network Time Protocol）サービスについて説明します。
第 7 章　システム構築：OS 環境構築	この章では、OS 環境構築の方法について説明します。
第 8 章　運用／保守：バックアップ／リストア	この章では、資産を確実に保護するために必要なバックアップ／リストアについて説明します。
用語解説	本文中で使用している専門用語について説明しています。

本書の表記

■ 本文中の記号

本文中に記載されている記号には、以下のような意味があります。

注 意	お使いになるときの注意点や、してはいけないことを記述しています。必ずお読みください。
参 考	理解を助けるための補足情報などを記述します。必要に応じてお読みください。
（→ P.xx）	参照先のページを示しています。クリックすると該当ページへ移動します。

■ 他のマニュアルの表記

本書以外の参照マニュアルを『(マニュアル名称)』と表記しています。
本文中では、本書以外に以下のマニュアルを参照しています。

● PRIMEQUEST 1000 シリーズの参照マニュアルの名称と表記

マニュアル名称（マニュアルコード）	本文中の表記
『PRIMEQUEST 1000 シリーズ 導入マニュアル』（C122-E107）	『導入マニュアル』
『PRIMEQUEST 1000 シリーズ 運用管理マニュアル』（C122-E108）	『運用管理マニュアル』
『PRIMEQUEST (1000 シリーズ) RHEL5-Xen 仮想マシン機能ユーザーズマニュアル』（J2UL-1200）	『RHEL5-Xen 仮想マシン機能ユーザーズマニュアル』
『PRIMEQUEST 1000 シリーズ VMware ゲスト OS インストールガイド Red Hat Enterprise Linux 5』（J2UL-1240）	『VMware ゲスト OS インストールガイド』
『PRIMEQUEST 1000 シリーズ システム構成図』 http://jp.fujitsu.com/platform/server/primequest/catalog/	『システム構成図』

● Red Hat 社から提供される公開マニュアル

マニュアル名称	本文中の表記
『Red Hat Enterprise Linux 5 Release Note』	『Release Note』
『Red Hat Enterprise Linux 5 Deployment Guide』	『Deployment Guide』
『Red Hat Enterprise Linux 5 Installation Guide』	『Installation Guide』
『Red Hat Enterprise Linux 5 DM Multipath』	『DM Multipath』
『Red Hat Enterprise Linux 5 Virtualization Guide』	『Virtualization Guide』

参 考

- **Red Hat 社から提供される公開マニュアルの入手方法については、**「1.1 Red Hat 社の情報活用」(→ P.13)**を参照してください。**

■ 製品表記

本文中では、製品名称を以下のように表記しています。

ソフトウェア製品名称	本書での表記	
Red Hat Enterprise Linux	RHEL	
Red Hat Enterprise Linux 5 (for Intel64)	RHEL5(Intel64)	RHEL5
Red Hat Enterprise Linux 5 (for x86)	RHEL5(x86)	
Red Hat Enterprise Linux 5.*n* (for Intel64)	RHEL5.*n*(Intel64)	RHEL5.*n* [注 1]
Red Hat Enterprise Linux 5.*n* (for x86)	RHEL5.*n*(x86)	

[注 1]："*n*" には Update 版数、またはマイナーリリースを示す数字が入ります。

ハードウェア製品名称	本書での表記	
FUJITSU Server PRIMERGY シリーズ	PRIMERGY シリーズ	PRIMERGY
FUJITSU Server PRIMEQUEST 1000 シリーズ	PRIMEQUEST 1000 シリーズ	PRIMEQUEST

■ コマンド入力

本文中では、コマンド入力を以下のように表記しています。

ユーザー可変（ユーザーの環境により異なる）の文字列

以下のように、斜体で表記します。

```
#/sbin/e2label <device> <label>
```

追加／変更対象文字列

以下のように、太字で表記します。

```
NETWORKING=yes
HOSTNAME=xxxx
・
・
VLAN=yes
```

説明を有する文字列

以下のように、文字列の下に下線を引き、枠下に説明文を記載しています。

```
disk = ["phy:/dev/sdd,hda,w"]
```

■ 容量の単位

データ容量を表す単位（KB,MB,GB）として、本書では断りのない限り、1KB=1024bytes 換算で表記しています。
ただし、ディスクサイズは一般的に 1KB=1000bytes 換算で扱われます。このため、ディスク領域を確保する操作においては、データ容量の見積り値をディスクサイズに換算してください。

■ キーの表記と操作方法

本文中のキーの表記は、キーボードに書かれているすべての文字を記述するのではなく、説明に必要な文字を以下のように記述しています。
例：【Ctrl】キー、【Enter】キー、【→】キーなど
また、複数のキーを同時に押す場合には、以下のように「＋」でつないで表記しています。
例：【Ctrl】+【F3】キー、【Shift】+【↑】キーなど

保守サービスについて

富士通では、お客様に安心して Linux を使用していただくために、有償サポート・サービス（SupportDesk Standard）をご用意しております。この有償サポート・サービスには、Linux に関するご質問、インストールや運用の際に発生する疑問やトラブルなどについて、富士通サポートセンター（OSC：One-stop Solution Center）にて一括対応させていただいております。また、Red Hat Enterprise Linux を利用する際に必要な、Red Hat 社からインストールイメージ（ISO ファイル形式）やアップデートを入手するためのサブスクリプション（利用権）が含まれています。有償サポート・サービスのご契約をお勧めします。
有償サポート・サービス（SupportDesk Standard）については、以下を参照してください。
http://jp.fujitsu.com/solutions/support/sdk/sd-standard/linux/

商標および著作権について

改版履歴

版数	変更日	変更箇所	修正概要
初版	2010 年 3 月 15 日	全章	初版作成
2 版	―	―	―
3 版	2010 年 5 月 31 日	第 6 章など	用語の変更、NTP の記事変更
4 版	2011 年 2 月 4 日	第 1 章	Red Hat 社公開情報を更新
		第 2 章	・ディスク容量見積り値を更新 ・DM-MP の注意点を追加
		第 3 章	mii 監視モードの情報を追加
		第 4 章	udev 機能の注意点を追加
		第 5 章	個別設定ファイルの注意点を追加
		第 7 章	・ゲスト OS（KVM）情報を追加 ・レスキューモードの手順を更新
		第 8 章	リモートバックアップの手順を追加
		全章	誤字・表記の誤りを修正
5 版	2011 年 8 月 11 日	第 2 章	誤記の修正、および補足の追加
		第 4 章	シングルユーザーモードの情報追加
		第 6 章	表 6-1 について記載内容の修正
		第 7 章	RHEL5.7 対応
6 版	2012 年 3 月 9 日	第 1 章、第 3 章、 第 4 章、 第 6 章～第 8 章	RHEL5.8 対応
7 版	2013 年 1 月 25 日	第 2 章、第 4 章、 第 6 章、第 8 章	RHEL5.9 対応
8 版	2013 年 10 月 15 日	第 3 章、第 7 章、 第 8 章	RHEL5.10 対応

目　次

第 4 章　システム設計：デバイス名ずれ対策設計

第 5 章　システム設計：ログファイル運用

第 6 章　システム設計：時刻補正の導入（NTP）

第 7 章　システム構築：OS 環境構築

1

第 1 章

概要：Linux のノウハウ

この章では、Linux をお使いになるうえで知っておいていただきたい情報と、Linux ご使用の流れについて説明します。

1.1 Red Hat 社の情報活用

Linux をお使いになるための様々な情報は、Red Hat 社から提供されています。
まずは Red Hat 社から提供されている以下の情報を確認し、活用してください。

- 「1.1.1 カスタマーポータル（Red Hat 社 Web サイト）」（→ P.13）
- 「1.1.2 オンラインマニュアル（man コマンド）」（→ P.14）

1.1.1 カスタマーポータル（Red Hat 社 Web サイト）

Red Hat 社のカスタマーポータル（https://access.redhat.com/home）では、RHEL を使用する際に有益な、様々なドキュメントを公開しています。

■ 製品マニュアル

Red Hat 社のカスタマーポータル（https://access.redhat.com/home）のリファレンスライブラリにおいて、RHEL5 の製品マニュアル（英語版）が公開されています。RHEL5 の基本的な利用方法に関するマニュアルは表 1 - 1 のとおりです。

表 1 - 1：公開ドキュメント

ドキュメント名称	参照のタイミング	内容
Release Note	導入／構築	マイナーリリースに関する新機能および更新情報などが記載されています。メジャーバージョンの初版では、このバージョン特有の機能について記載されています。
Deployment Guide	導入／構築	システムの設定および、カスタマイズ方法が記載されています。お読みになるには、システムの基礎知識が必要です。
Installation Guide	導入／構築	一般的なインストールについての関連情報が記載されています。[注 1]
DM Multipath	導入／構築／運用・保守	DM-MP（Device-Mapper Multipath）機能を利用するための情報が記載されています。
Virtualization Guide	導入／構築／運用・保守	RHEL5 が提供する仮想機能の設定・管理・トラブルシューティング情報が記載されています。

［注 1］：PRIMEQUEST 1000 シリーズで Linux をお使いになる場合、富士通が提供する独自のインストーラ（ServerView Installation Manager（以降 SVIM と表記します））が添付されています。SVIM を使用する場合は、Red Hat 社の『Installation Guide』に記載されているインストール方法は参照しません。

製品マニュアルの日本語対応については、日本語向けのカスタマーポータルを確認してください。本書は、最新の情報が反映されるマニュアルとして、英語圏向けのカスタマーポータルで公開されるマニュアルを参照対象とします。

参 考

- **RHEL5 の製品マニュアル（英語版）を参照するには、カスタマーポータル画面の最下層にある言語の選択で「English」を選択し、英語圏向けのカスタマーポータルに表示を切り替えてください。英語圏向け以外のカスタマーポータルでは、RHEL5 の製品マニュアルが配置されていない場合があります。**

■ ナレッジベース

Red Hat 社のカスタマーポータル（https://access.redhat.com/home）のナレッジベースでは、利用者からよくある質問と回答を公開しています。トラブルが発生した場合や、困ったときなど、参考情報があるかどうかを確認してください。

注 意

- **ナレッジベースを利用の際は、事前に Red Hat 社製品のサブスクリプション登録が必要です。サブスクリプション登録については、****「7.1 カスタマーポータルへのサブスクリプション登録」（→ P.150）****を参照してください。**

1.1.2 オンラインマニュアル（man コマンド）

man コマンドは、オンラインマニュアルを参照するコマンドです。man コマンドを使用して、コマンド・関数・設定ファイルなどの使用方法を確認できます。
ここでは、基本的な使い方を紹介します。man コマンドの詳細説明については、man コマンドのオンラインマニュアルを参照してください。

■ 書式

```
man ［オプション］［セクション］ 名前 または キーワード
```

名前には、コマンド名、関数名、設定ファイル名などを入力します。

■ セクション

オンラインマニュアルは、以下のセクション（章）に分かれています。

表 1 - 2：オンラインマニュアルのセクション

章番号	内容
1	実行プログラムまたはシェルのコマンド
2	システムコール（カーネルが提供する関数）
3	ライブラリコール（システムライブラリに含まれる関数）
4	スペシャルファイル（通常 /dev に置かれている）
5	ファイルのフォーマットとその規約
6	ゲーム
7	マクロのパッケージとその規約
8	システム管理用のコマンド
9	カーネルルーチン

■ マニュアルの構成

オンラインマニュアルは、以下の構成で記述されています。

表 1 - 3：オンラインマニュアルの構成

項目	内容
名前（NAME）	マニュアルページの対象の名前です。通常はコマンドかファイル名と、それについての簡単な説明文が記載されています。
書式（SYNOPSIS）	コマンドを呼び出す際に使用される構文の概要が記載されています。
説明（DESCRIPTION）	使い方やその他全般的な理解に必要な詳細情報について記載されています。
オプション（OPTION）	コマンドを実行する際に利用できるオプションが記載されています。
ファイル（FILE）	マニュアルページのトピックに関係のあるファイル（例えば設定ファイルなど）について記載されています。
関連項目（SEE ALSO）	他の関連情報が列挙されています。
バグ（BUGS）	マニュアルページのソフトウェアに存在する既知のバグが記載されています。
作者（AUTHOR）／履歴（HISTORY）	マニュアルページのソフトウェアを誰が書いたか、または誰がその開発に貢献したかの情報が記載されています。

■ 使用例

コマンド名がわかっている場合

```
$ man ls

LS(1)                                                              LS(1)

名前
       ls, dir, vdir - ディレクトリの中身をリスト表示する

書式
       ls [options] [file...]

       POSIX オプション: [-CFRacdilqrtu1]

       GNU  オプション (簡略形式): [-1abcdfghiklmnopqrstuvxABCDFGHLNQRSUX] [-w
        cols]  [-T  cols]  [-I  pattern]  [--block-size=SIZE]  [--classify]
       [--color[={none,always,auto}]]  [--file-type]  [--full-time]  [-for
       -mat={across,commas,long,single-column,verbose,vertical}] [--human-
       read-able]       [--indicator-style={none,file-type,classify}]
       [--quot-ing-style={c,clocale,escape,literal,locale,shell,shell-
       always}]  [--show-control-chars]  [--si]  [--sort={none,extension,
       size,time,ver-sion}]  [--time={atime,access,ctime,status,use}]  [--
       help]  [--version]  [--]

説明
       プログラム ls は、最初にディレクトリでない引数 file をリスト表示する・・・
```

コマンド名がわからない場合

キーワード検索でキーワードが含まれるマニュアルを検索します。

```
$ man -k passwd

chpasswd             (8)  - update password file in batch
ckpasswd             (1)  - nnrpd password authenticator
gpasswd              (1)  - administer the /etc/group file
hesiod_free_passwd [hesiod_getpwnam] (3)  - Hesiod functions for retrieving
passwd information
hesiod_getpwnam      (3)  - Hesiod functions for retrieving passwd
information
hesiod_getpwuid [hesiod_getpwnam] (3)  - Hesiod functions for retrieving
passwd information
htpasswd             (1)  - Manage user files for basic authentication
ldappasswd           (1)  - change the password of an LDAP entry
lppasswd             (1)  - add, change, or delete digest passwords
mkpasswd             (1)  - generate new password, optionally apply it to a
user
pam_localuser        (8)  - require users to be listed in /etc/passwd
pam_passwdqc         (8)  - Password quality-control PAM module
passwd               (1)  - update a user's authentication tokens(s)
passwd               (5)  - password file
• • •
```

1.2 Linux ご使用の流れ

Linux をご使用になる場合、システム設計から運用／保守までに必要な作業は以下のとおりです。

「1.2.1 システム設計」（→ P.17）
- **ハードディスク設計**
- **ネットワーク設計**
- **デバイス名ずれ対策設計**
- **ログファイル運用設計**
- **時刻補正の導入**

⬇

「1.2.2 システム構築」（→ P.18）
- **OS のインストール**

⬇

「1.2.3 運用と保守」（→ P.18）
- **バックアップ／リストア**

1.2.1 システム設計

■ ハードディスク設計

OS をインストールするために必要なパーティション設計や、ハードディスク冗長化など、ハードディスクに関する構成設計を行います。
詳しくは、「第 2 章 システム設計：ハードディスク」（→ P.19）を参照してください。

■ ネットワーク設計

サーバが使用する LAN には、業務 LAN、管理 LAN、プライベート LAN の 3 種類があります。ネットワーク設計では、業務 LAN において、VLAN や bonding ドライバによる二重化、ロードバランシングの設計を行います。
詳しくは、「第 3 章 システム設計：ネットワーク」（→ P.29）を参照してください。

■ デバイス名ずれ対策設計

起動時のデバイスの認識順序が前回起動したときと異なる場合に「デバイス名ずれ」が発生します。これにより、システムが起動できないトラブルや、データ参照ができないトラブルが想定されます。デバイス名ずれに対応するために必要な設計を行います。
詳しくは、「第 4 章 システム設計：デバイス名ずれ対策設計」（→ P.46）を参照してください。

■ ログファイル運用設計

システムやアプリケーションで発生した様々な事象は、システムやアプリケーションが作成するログファイルに記録されます。ログファイル運用設計では、システムログファイルのローテーションの設計を行います。
詳しくは、「第 5 章 システム設計：ログファイル運用」（→ P.89）を参照してください。

■ 時刻補正の導入

時刻補正とは、ハードウェア内部の時計の時刻を補正するための設定です。運用を開始する前に、時刻補正の導入を行ってください。
詳しくは、「第 6 章 システム設計：時刻補正の導入（NTP）」（→ P.110）を参照してください。

1.2.2 システム構築

■ OS のインストール

Linux のインストール方法は、ハードウェアの機種により異なります。
詳しくは、「7.2 RHEL のインストール」（→ P.151）を参照してください。

1.2.3 運用と保守

■ バックアップ／リストア

お客様資産を確実に保護するためには、ファイルシステムのバックアップとリストアを対にして、計画的な運用を立案し実施してください。
詳しくは、「第 8 章 運用／保守：バックアップ／リストア」（→ P.157）を参照してください。

2

第 2 章

システム設計：ハードディスク

この章では、ハードディスク関連の設計情報を説明します。

2.1 ディスクパーティションの容量見積り

ディスクパーティションの構成は、各システムの用途によって異なります。

• 「2.1.1 見積り例」（→ P.20）

ここでは OS の動作に必要な以下のパーティションについて説明します。

- /（root パーティション）
- /boot パーティション
- swap パーティション

ミドルウェアなどをインストールする場合、およびメモリダンプ機能の設定を行う場合は、領域の確保が別途必要となります。それぞれの領域要件を考慮し、パーティション構成を決定してください。

2.1.1 見積り例

各ディスクパーティションの容量を見積る際は、各パーティションの用途、容量の目安を考慮します。

ディスクの容量は、1KB=1024byte で計算してください。

例：73GB の容量確保が必要な場合

73 × 1024 × 1024 × 1024 ＝ 78,383,153,152byte の領域が必要です。

これに対し、ディスク容量は通常 1KB=1000byte で表記されます。このため、「73GB ディスク」では、実際の容量は 73,000,000,000byte となり、必要な容量に達しません。この例においては、78.4GB 以上のディスクを用意する必要があります。

また、パーティション領域の数 % がファイルシステムの管理領域として使用されます。容量は余裕を持って見積ってください。

■ 各パーティションの容量見積り

表 2 - 1：各パーティションの見積り（1KB=1024byte 換算）

パーティション	用途	推奨容量
/	root ディレクトリ	20GB 以上
/boot	ブートストラップのプロセス中に使用されるファイルとともに、OS のカーネルが格納される領域	256MB
swap	現在動作中のプロセスを取り扱うだけの十分なシステムメモリがないとき、仮想記憶域として使用される領域	システムメモリの容量に応じて、推奨容量が異なります。 メモリ容量ごとの推奨容量については、「● swap 見積りの考慮」（→ P.21）を参照してください。

■ 各パーティションの容量見積りの考慮

● /（root パーティション）見積りの考慮

20GB 以上の容量を確保してください。

- 機能追加／障害修正の適用
  OS インストール後に、以下の機能追加／障害修正を行うことで、ファイルシステムの使用量が増加する可能性があります。
  - ISV のソフトウェア、オプションソフトウェアの追加
  - OS のソフトウェアグループの変更
  - パッチの適用など

  インストール時には、余裕を持ってファイルシステムサイズを設定してください。
  /usr パーティション、/opt パーティションを設定している場合も同様です。
- ログファイルの保存
  var ディレクトリには、メールボックスやプリンタスプールの他に、OS や各種ミドルウェアのログファイルが保存されます。
  ログファイルは、保存期間によって予期せぬサイズとなることがあります。十分なサイズを設定してください。

注 意

- **SVIM で「すべてインストール」を選択した場合、10GB ではインストール時に容量不足が発生する可能性があるため、20GB 以上を推奨します。**

● swap 見積りの考慮

swap 領域の容量は、システムにかかるメモリ負荷に応じて設定する必要があります。
しかし、システム導入時にメモリ負荷を見積もることが難しい場合は、コンピュータの RAM（物理メモリ）容量に応じて、表 2 - 2 を参考に設定してください。

表 2 - 2：swap パーティションのサイズ（1KB=1024byte 換算）

RAM 容量（GB）	最低限必要な swap 領域の容量（GB）
M ＜ 4	2GB
4 ≦ M ＜ 16	4GB
16 ≦ M ＜ 64	8GB
64 ≦ M ＜ 256	16GB
256 ≦ M ＜ 512	32GB

M：RAM（物理メモリ）容量（GB）

参 考

- **swap の設計に関する詳細については、Red Hat 社のマニュアル『Installation Guide』を参照してください。**

2.2 ハードディスクの冗長化

LVM や MD ／ DM-MP を利用して、ハードディスクを管理することで、フレキシブルなストレージ管理を可能にしたり、データの可用性と信頼性を向上させたりすることができます。

- 「2.2.1 LVM」（→ P.22）
- 「2.2.2 MD」（→ P.24）
- 「2.2.3 DM-MP」（→ P.25）

2.2.1 LVM

LVM は、通常のファイルシステムやパーティションよりも高度なストレージを提供し、管理者により一層フレキシブルなストレージ管理を可能にします。
LVM の大きな特徴は以下のとおりです。

- 複数のドライブからなる物理的な記憶装置をまとめ、論理的な記憶装置（論理ボリューム）を作成することができる。
- 論理ボリュームは必要に応じて記憶装置を追加・削除することができる。
- 論理ボリューム上に作成したパーティションは拡張・縮小することができる。
- 論理ボリューム上に作成したパーティションのスナップショットを作成することができる。

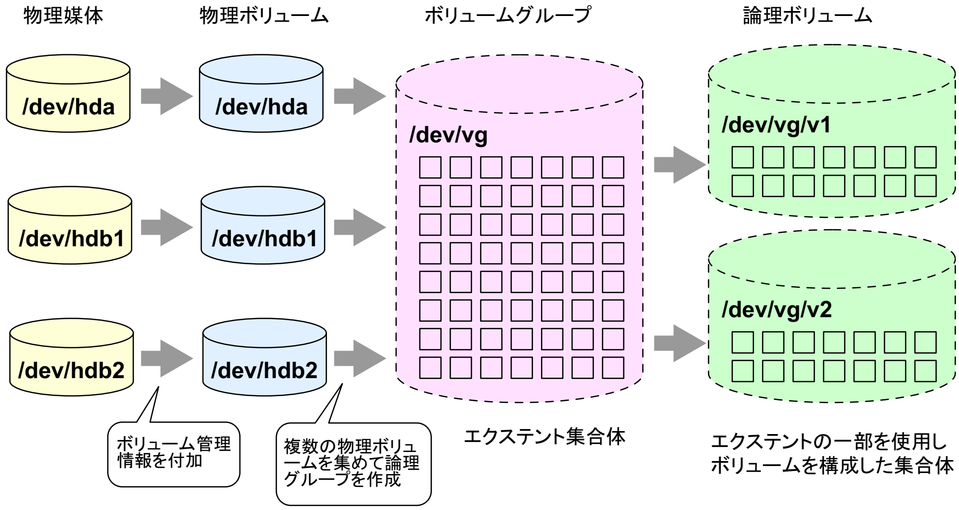


図 2 - 1：LVM のイメージ

■ LVM 構成要素

LVM を扱うには、まずその構成要素の名称と役割を理解する必要があります。理解しないまま変更などを加えると、ファイルシステムを破壊するなどの危険性があるので注意してください。

表 2 - 3：LVM 構成要素

構成要素	略称	説明
物理媒体	─	通常のハードディスクやパーティション（例：/dev/sda、/dev/sda1）。 パーティションを LVM の物理媒体として使用する場合には、パーティションタイプを「Linux LVM (8e)」とする必要があります。
物理ボリューム	PV	ボリューム管理情報を付加した物理媒体。 物理ボリュームを作成することで初めて LVM で使用可能な領域（エクステント）を確保できます。
物理エクステント	PE	物理ボリュームをある一定サイズで区切ったデータ単位。
論理エクステント	LE	物理エクステントとまったく同じもの。 物理ボリューム上のエクステントを物理エクステント、論理ボリューム上のエクステントを論理エクステントと呼びます。機能的な差分はありません。
ボリューム・グループ	VG	複数の物理ボリュームを集めて作成する論理グループ。 この中から実際にパーティションとして使用する論理ボリュームを切り出すことができます。エクステントのサイズは、ボリューム・グループ作成時に指定します。
論理ボリューム	LV	ボリューム・グループよりいくつかのエクステントを集めて作成する論理的なパーティション。 通常のファイルシステムと同様にフォーマット／マウントして利用できます。

■ LVM ボリューム状態

表 2 - 4：LVM ボリューム状態

ボリューム状態	説明
アクティブ	ボリューム情報がカーネル空間に存在し、実際のストレージとして利用可能な状態を指します。
インアクティブ	ボリューム情報がユーザーランド（ユーザー空間）に存在し、論理的なボリュームの新規作成や修正作業が可能な状態を指します。 LVM 関連コマンドは一部を除き、アクティブ／インアクティブの状態を自動で変更します。

2.2.2 MD

MD はドライバ層で RAID を実現するソフトウェア RAID です。
MD を利用した RAID 機能では、複数台のハードディスクを組み合わせて、ディスク故障時にデータの損失防止やディスクアクセスの処理速度の向上を実現します。
MD では表 2 - 5 の RAID レベルをサポートします。

表 2 - 5：MD でサポートする RAID レベル

RAID レベル	機能
linear	複数台のディスクを使用して、1 つの大容量の仮想ドライブを作る。
RAID0	ストライピングとも呼ばれ、複数台のディスクにデータを分割して書き込む。IO 性能は向上するが、冗長性がない。
RAID1	ミラーリングとも呼ばれ、複数台のディスクに同じデータを書き込む。IO 性能は向上しないが、冗長性がある。
RAID10	RAID0 と RAID1 を組み合わせたもの。IO 性能は向上し、冗長性がある。

参 考

- **RAID レベルの詳細については、以下の「用語解説 RAID」を参照してください。**
  http://storage-system.fujitsu.com/jp/term/raid/

構成としてはデータボリューム域の RAID 構成をサポートします。
データボリューム域の RAID 構成は、OS のインストール後に mdadm コマンドで構成します。
図 2 - 2 では、/home を RAID1 で構成し、同じサイズのパーティションをディスク 2、ディスク 3 に配置しています。

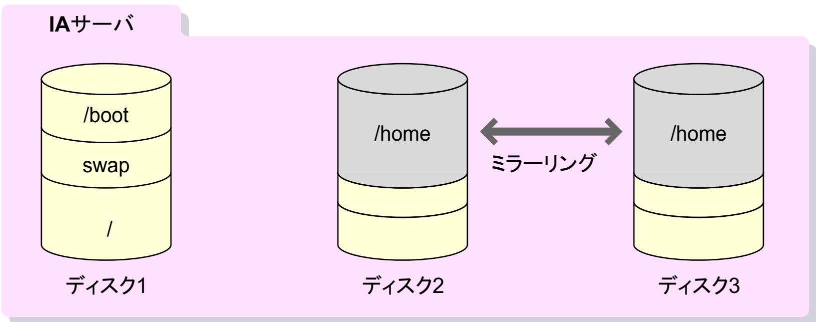


図 2 - 2：MD の構成

2.2.3 DM-MP

DM-MP を利用したマルチパス機能では、冗長性を持たせたディスクアレイ装置へのアクセスパスを統合して、アクセスパス障害の発生に備えたシステム運用の継続性を実現します。

参 考

- **DM-MP の詳細については、Red Hat 社のマニュアル『DM Multipath』を参照してください。**

注 意

- **富士通製「ETERNUS マルチパスドライバ」など、DM-MP と同等の機能を提供する製品をお使いの場合は、DM-MP 機能は併用しないでください。**

SAN 環境において DM-MP を利用することで、ディスクデバイスへのアクセスパスを簡単にマルチパスで構成できます。DM-MP は、以下のどちらの形態もサポートしています。経路となるアクセスパスは複数本で構成しておく必要があります。

- ファイバーチャネルスイッチを経由する SAN 構成

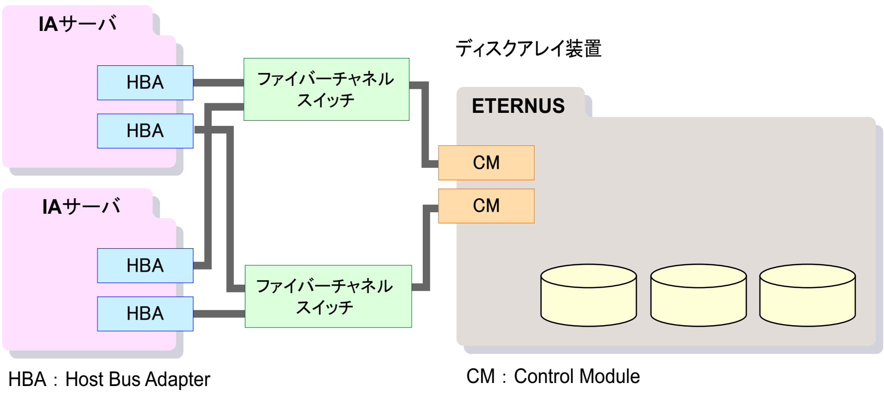


図 2 - 3：ファイバーチャネルスイッチを経由する SAN 構成

- ファイバーチャネルケーブルをサーバに直結する構成

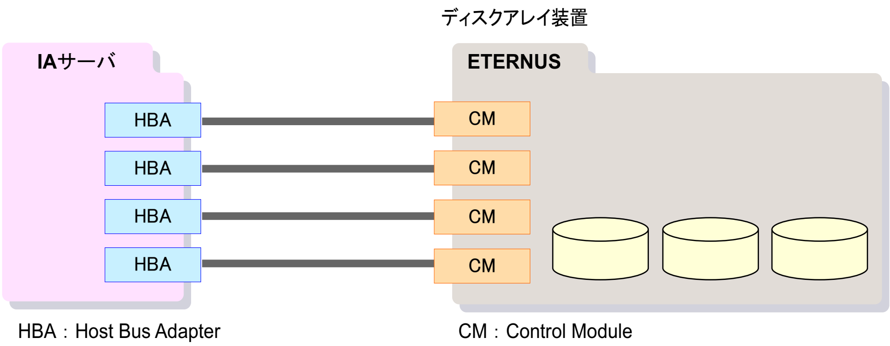


図 2 - 4：ファイバーチャネルケーブルを直結する構成

マルチパスのディスク構成としては、以下の構成があります。

- ディスクアレイ装置にシステムボリュームを配置した SAN Boot のマルチパス構成
「■ SAN Boot（システムボリューム）のマルチパス構成」（→ P.26）
- ディスクアレイ装置にデータボリュームを配置したデータボリュームのマルチパス構成
「■ データボリュームのマルチパス構成」（→ P.27）

マルチパスの方式としては、以下の 2 つをサポートしています。

- 「■ 運用待機方式（failover）」（→ P.27）
- 「■ 負荷分散方式（multibus）」（→ P.28）

■ マルチパス構成のデバイスファイル

マルチパス構成にしたディスクデバイスは、表 2 - 6 のデバイスファイル名でアクセスできます。

表 2 - 6：マルチパス構成のデバイスファイル

種別	デバイスファイル名	例：alias に mpath0 を指定した場合
ボリューム	/dev/mapper/ ＜*NNNNN* ＞ ［注 1］	/dev/mapper/*mpath0*
パーティション	/dev/mapper/ ＜*NNNNN* ＞pM ［注 1］［注 2］	/dev/mapper/*mpath0p1*

［注 1］：NNNNN はボリュームを表す名前（alias 名）を示します。
［注 2］：pM （文字列 'p' + 数字）はパーティション番号を示します。数字は自動的に割り振られます。

■ SAN Boot（システムボリューム）のマルチパス構成

システムボリュームをディスクアレイ装置上に配置する SAN Boot の構成をマルチパスで構成できます。

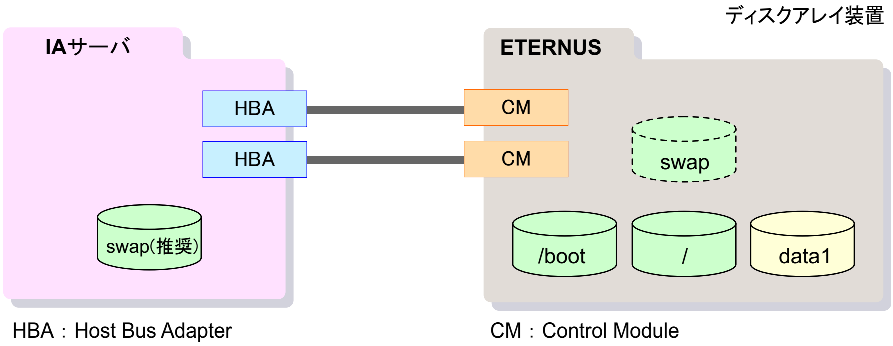


図 2 - 5：SAN Boot（システムボリューム）のマルチパス構成

注 意

- **swap パーティションは頻繁にアクセスするため、内蔵ハードディスクに配置することをお勧めします。**
- **/boot、/ を含むシステムボリュームはディスクアレイ上に配置してください。**

■ データボリュームのマルチパス構成

ディスクアレイ装置上にデータボリュームを配置して、マルチパスで構成することができます。

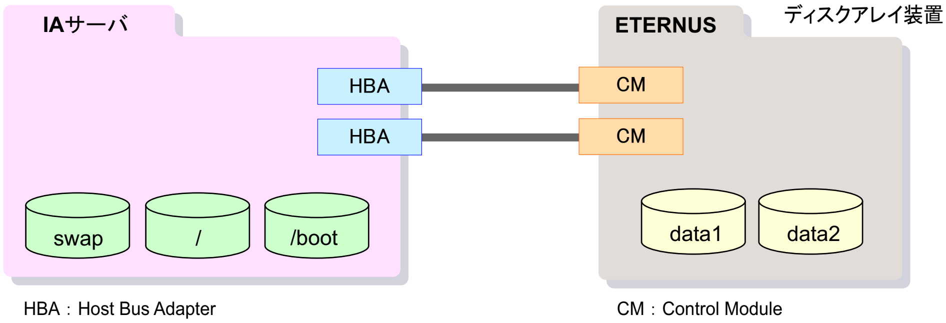


図 2 - 6：データボリュームのマルチパス構成

■ 運用待機方式（failover）

アクセスパスの 1 つを運用系、残りを待機系として運用する方式です。

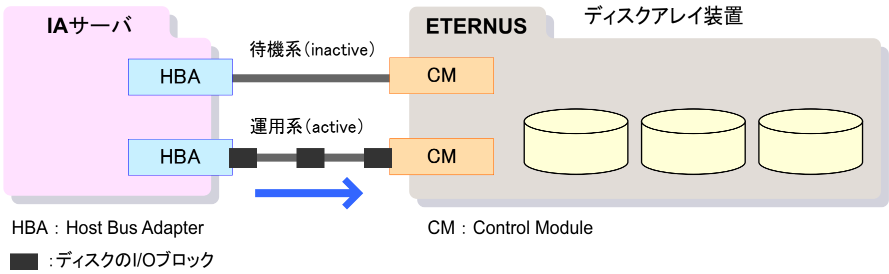


図 2 - 7：運用待機方式（failover）

図 2 - 7 は、運用系（active）と待機系（inactive）のアクセスパスでのディスク I/O の流れを示しています。
運用系パスで故障が発生すると、待機系パスの 1 つが自動的に運用系（active）へ切り替わります。

■ 負荷分散方式（multibus）

すべてのアクセスパスを使用して、ディスク I/O のブロックをラウンドロビンで負荷分散することで、アクセス性能を向上させる方式です。

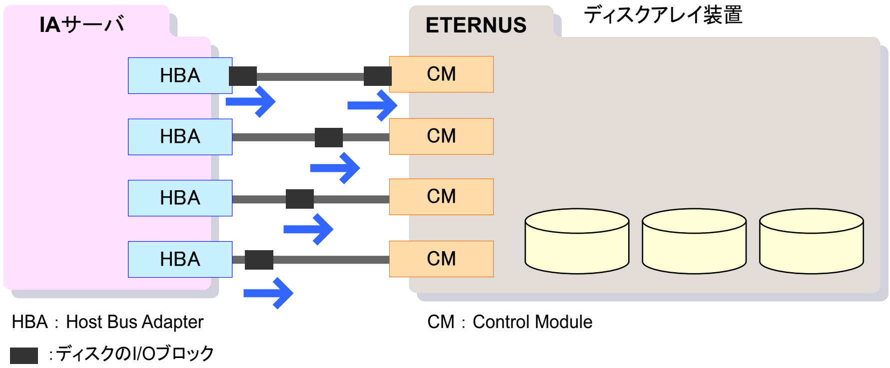


図 2 - 8**：負荷分散方式**（multibus）

図 2 - 8 は、4 つのパスすべてを使用して負荷分散した際のディスク I/O の流れ（ラウンドロビン）を示しています。
アクセスパスで故障が発生すると、残りのアクセスパスで運用を継続します。

3

第 3 章

システム設計：ネットワーク

この章では、ネットワークの設計に関する設定方法や注意事項について説明します。

3.1 LAN の構成

サーバ本体が接続する LAN の構成について説明します。

- 「3.1.1 LAN 構成の考慮点」（→ P.30）
- 「3.1.2 Linux における LAN 接続機能」（→ P.31）

3.1.1 LAN 構成の考慮点

サーバ本体が接続する LAN を構成する際の考慮点について説明します。

筐体内に複数のサーバ装置を収納するハードウェアでは、各パーティションを管理するための装置（マネージメントボード）が搭載されています。
マネージメントボードは、ネットワークを介してシステム管理機能を遠隔地の管理コンソールや管理ソフトウェアに提供します。このため、セキュリティと負荷分散を目的として、独立した以下の 2 種類の LAN を構成することを推奨します。

- 業務 LAN
  お客様の業務システムを構築するための LAN です。
- 管理 LAN
  システムを管理するために用いる LAN です。

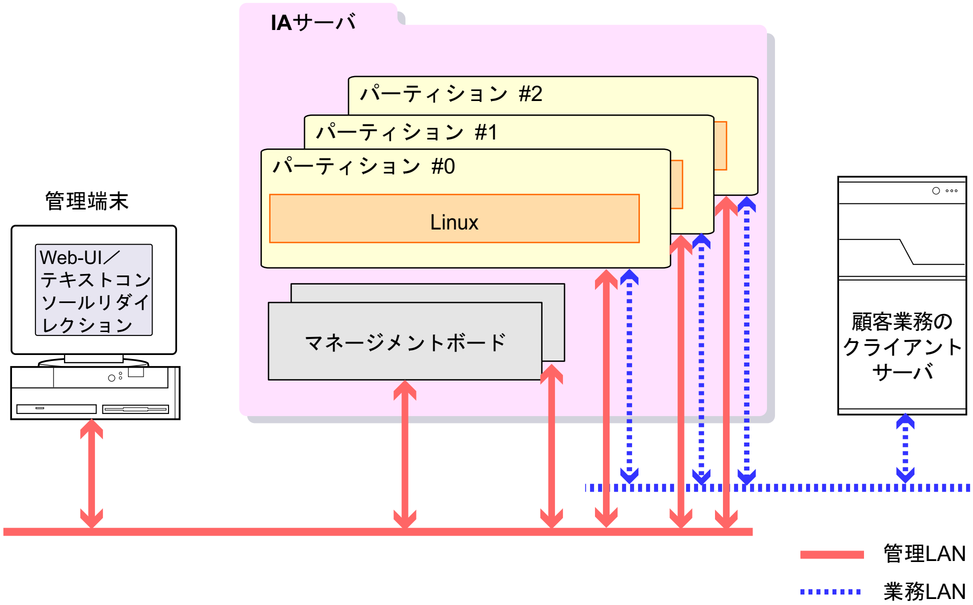


図 3 - 1：管理装置を搭載する機種の LAN 構成例

3.1.2 Linux における LAN 接続機能

LAN の構成設計は、ネットワーク全体の構成設計の一部分であり、Linux 利用技術の範囲を超えるため、本書での解説は割愛します。以降では、システムを LAN に接続するにあたって利用可能な Linux の機能について説明します。

- VLAN
  物理的な 1 つの伝送路上に、仮想的なネットワークグループ（VLAN グループ）が設定された LAN 環境に Linux システムを接続する場合の設定方法および注意事項について説明します。詳細は、「3.2 VLAN」（→ P.32）を参照してください。
- bonding ドライバ
  Linux は、サーバ本体に搭載された複数の NIC を、仮想的な 1 つの NIC として扱う機能を有します。これを bonding 機能と呼びます（チーミング機能とも呼びます）。この bonding 機能の用途と、設定方法および注意事項について説明します。詳細は、「3.3 bonding ドライバ」（→ P.37）を参照してください。

3.2 VLAN

VLAN **ネットワークを構築する場合の設定方法や、注意事項について説明します。**

- 「3.2.1 VLAN **とは」（→** P.32）
- 「3.2.2 VLAN **ドライバの設定」（→** P.34）
- 「3.2.3 VLAN **使用時の注意事項」（→** P.34）

3.2.1 VLAN とは

VLAN は、物理的な 1 つの伝送路を、複数の仮想的なネットワークのグループ（VLAN グループ）に分割する機能です。同じ伝送路上の機器同士でも、VLAN グループが異なると通信ができません。VLAN を使用すると、以下のような効果があります。

- 不必要なパケットが異なる VLAN グループへは送信されないため、ネットワークトラフィックを抑制できる。
- ネットワークをグループ化することにより、ネットワーク資源への不要なアクセスを制限できる。

VLAN をサポートするスイッチは、portVLAN と、tagVLAN の 2 種類のインターフェースを提供します。

■ portVLAN

portVLAN とは、VLAN を実現する方式の 1 つです。
LAN スイッチのポートごとに VLAN 番号を設定する方法で、スイッチ内で同じ VLAN 番号を持つポートが、1 つのセグメントとして定義されます。これにより、スイッチ内を論理的に分割することができます。
Linux システムを portVLAN に接続する場合、特別な設定は必要ありません。

ポートとVLANの対応表

ポート	VLAN番号
1	1
2	1
3	1
4	1
5	2
6	2
7	2
8	2

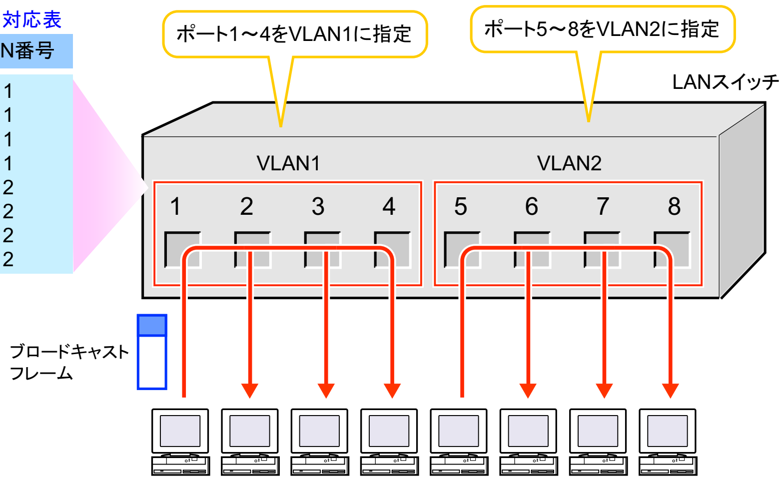


図 3 - 2**：**portVLAN **の概要**

■ tagVLAN

tagVLAN とは、ネットワークを流れるフレームに「VLAN タグ」と呼ばれる VLAN 識別情報を付加することで、複数のセグメントを論理的に構築する VLAN の方式です。「VLAN タグ」には VLAN 番号が設定され、これにより VLAN の識別が行われます。tagVLAN は、1 つのポートを複数の VLAN に所属させる場合に用いられます。
tagVLAN を構築するには、スイッチやサーバ機のポートが「VLAN タグ」付きフレームの送受信を可能とするための、タグ付きポートとしての設定が必要になります。

tagVLAN は、主に以下の用途で用いられます。

- 複数のスイッチにまたいで VLAN を構成する場合のスイッチ間の接続
  複数のスイッチにまたいで VLAN を構成する場合、スイッチ間の接続を tagVLAN で構成することにより、物理的な接続が 1 つで済むようになります。
- サーバ機を 1 つの NIC で複数のネットワークに接続する場合
  ネットワークやネットワークに属する機器を監視するなど、業務上の目的でサーバ機を複数のネットワーク・セグメントに所属させる必要がある場合、サーバ機が tagVLAN に接続可能であれば、1 つの NIC で複数のネットワーク・セグメントへの通信が行えるようになります。

Linux システムでは、VLAN ドライバを用いることで「VLAN タグ」付きフレームを送受信することが可能となり、tagVLAN への接続が可能となります。設定方法の詳細は、「3.2.2 VLAN ドライバの設定」（→ P.34）を参照してください。

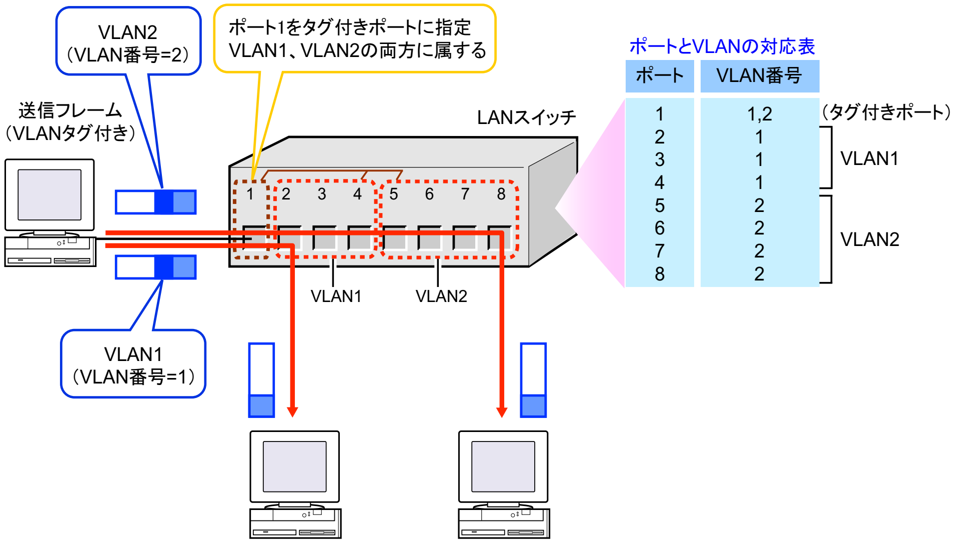


図 3 - 3：tagVLAN の概要

3.2.2 VLAN ドライバの設定

tagVLAN を使用する場合の VLAN ドライバの設定方法を示します。ここでは、設定ファイルによる方法を記述します。

注 意

- **bonding と組み合わせる場合の注意事項については、****「■ VLAN ネットワークとの組み合わせ時の注意」（→ P.43）****を参照してください。**

1 tagVLAN に相当する構成ファイルを追加します。

eth1 または bond0 上に VLANID「2」を割り当てた場合を例に説明します。

参 考

- **tagVLAN に相当する構成ファイルの形式は、「ifcfg-*xxxx.n*」です。**
  「*xxxx*」は tagVLAN を載せるインターフェース名、「*n*」は VLANID です。

eth1 に VLANID「2」を割り当てた場合

「/etc/sysconfig/network-scripts/ifcfg-eth1.2」を追加し、以下のように設定します。

```
DEVICE=eth1.2
BOOTPROTO=static
IPADDR=x.x.x.x
ONBOOT=yes
VLAN=yes
  •
```

bond0 に VLANID「2」を割り当てた場合

「/etc/sysconfig/network-scripts/ifcfg-bond0.2」を追加し、以下のように設定します。

```
DEVICE=bond0.2
BOOTPROTO=static
IPADDR=x.x.x.x
ONBOOT=yes
VLAN=yes
  •
```

3.2.3 VLAN 使用時の注意事項

VLAN を使用する場合、スイッチの持つ一部の機能が制限されることがあります。
注意事項について説明します。

■ STP の設定

tagVLAN、portVLAN を問わず、VLAN を導入したスイッチネットワーク で、STP を正しく動作させるには、拡張された STP（MSTP）が必要です。
VLAN を導入し、冗長化したスイッチネットワークが必要な場合には、MSTP をサポートしているスイッチを導入してください。

● STP が不要な例

図 3 - 4 で、Host A、Host C と Host B、Host D は独立したネットワークです。スイッチ 1 とスイッチ 2 を portVLAN（赤、青）で分割して使用しています。スイッチ間の渡りは tagVLAN（黄色）でトランキングしています。閉路がないため STP は不要です。

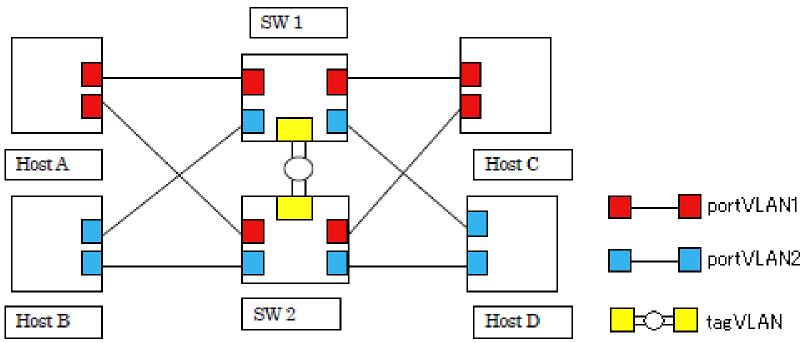


図 3 - 4：STP が不要なネットワーク

● MSTP が必要な例

図 3 - 5 は STP の不要な例にスイッチを 1 ペア追加し、VLAN ごとに異なる冗長路を作成した例です。閉路があるので STP は必須です。太線は、STP のみ（MSTP でない）の場合に有効とされるリンクの例で、どちらの VLAN にも到達できないスイッチが出現します。
MSTP では、VLAN を意識して太破線も有効リンクに追加されるので、到達できないスイッチは出現しません。
実際に、どのリンクが選ばれるかは、冗長性を保証する範囲で、タイミングなどにより決まります。図 3 - 5 の接続形態であれば、必ず図示したリンクが選ばれるとは限りません。

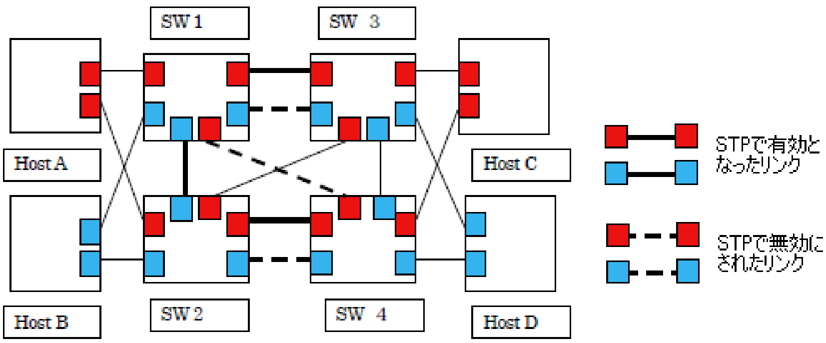


図 3 - 5：MSTP が必要なネットワーク

■ スイッチの IP アドレス

スイッチには、IP アドレスを付与することができますが、一般的にスイッチを複数の VLAN で分割すると、その中の 1 つの VLAN からしか、スイッチの IP アドレスへアクセスできません。
例えば、GLS を導入するとスイッチに監視用の IP アドレスが必要となります。このとき、スイッチを portVLAN の設定により、独立した複数の区画に分割して、それぞれの区画を、GLS に独立したスイッチとして認識させようとすると、片方の区画には監視用の IP アドレスが存在しないことになり、導入がうまくできません。
スイッチの IP アドレスを監視用途に使用する機能を導入する場合は、このことを考慮して設計を行ってください。
または、複数の VLAN で分割しても、すべての VLAN からスイッチの IP アドレスにアクセスできる機能（VLAN ごとに IP アドレスを割り当てることができる機能）を持ったスイッチを使用してください。

● IP と VLAN の関係

図 3 - 6 はスイッチ 1、スイッチ 2 とも portVLAN（赤、青）で分割して使用しています。IP1、IP2 とも赤の VLAN のポートに割り当てられていて、青のポートからはアクセスできません。

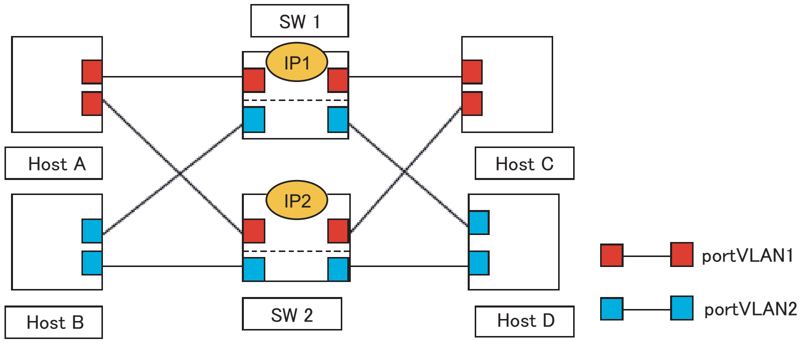


図 3 - 6：IP と VLAN の関係

3.3 bonding ドライバ

bonding ドライバを使用した二重化機能／ロードバランシング機能の設定方法や、注意事項について説明します。
bonding とは、複数の NIC を組み合わせて、仮想的な 1 つの NIC として扱う Linux の機能です。

3.3.1 bonding ドライバの動作モードと監視モード

bonding ドライバには 7 種類の動作モードと、2 種類の監視モードがあります。

■ 動作モード

bonding ドライバには 7 種類の動作モードがあり、それぞれ使用目的に応じて使い分けます。
bonding ドライバの動作モードの特徴は、表 3 - 1 のとおりです。

表 3 - 1：bonding ドライバの動作モード

動作モード	動作の特徴
mode0 (balance-rr)	ロードバランシング機能 スイッチのトランキングのサポートが必要 データの分散方式はラウンドロビン
mode1 (active-backup)	二重化機能専用 現用ー待機構成をとり、1 時点でアクティブなリンクは 1 本のみ
mode2 (balance-xor)	ロードバランシング機能 スイッチのトランキングのサポートが必要 データの分散方式は相手によるリンク固定
mode3 (broadcast)	ブロードキャスト用 本マニュアルでは説明の対象外
mode4 (802.3ad)	ロードバランシング機能 スイッチのトランキングのサポートが必要 データの分散方式は、IEEE802.3ad 準拠
mode5 (balance-tlb)	ロードバランシング機能 スイッチのサポートが不要な方式 データの分散方式は、負荷分散方式
mode6 (balance-alb)	ロードバランシング機能 スイッチのサポートが不要な方式 データの分散方式は、mode5 と同じだが、受信も分散可 IPv4 のみ適用可能

■ 監視モード

bonding ドライバには、各 NIC が通信可能な状態であるかを監視する機能が 2 種類あり、この種別を監視モードと呼びます。
監視の有無の指定、および監視モードの指定は、いずれも bonding ドライバのオプション設定により行います。
監視モードは以下のとおりです。

● mii 監視モード

NIC とネットワーク機器とのリンク状態を直接監視する方式です。ケーブル抜けやネットワークスイッチの故障などを直接検出します。

● arp 監視モード

arp 要求パケットをネットワークに送出し、それに対してネットワークからのフレーム受信の有無を監視する方式です。何らかのフレーム受信が検出できれば、通信可能な状態と判断します。

3.3.2 二重化機能

ネットワークを二重化し、一方のネットワークをバックアップ用として待機させておきます。運用中のネットワークで障害（NIC 故障、スイッチ故障、ケーブル切断など）が発生したときに、バックアップ用のネットワークに切り替えることで、ネットワークの停止を防止する方法です。

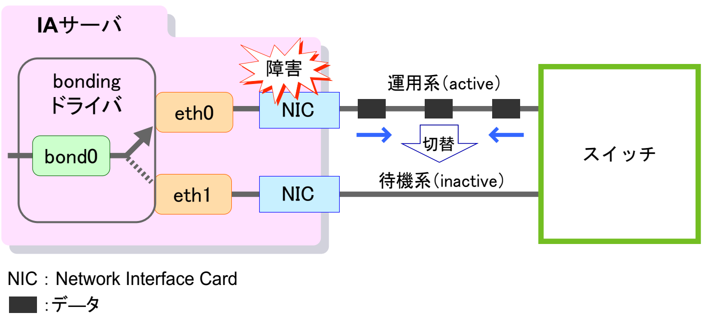


図 3 - 7：二重化機能の概要

ここでは、bonding ドライバを使用した LAN の二重化を構築する場合の推奨モード、設定手順を説明します。

■ 推奨する設定

- 動作モード
  二重化の場合、動作モードは mode1 のみとなります。
- オプションの設定
  mii 監視モードを推奨します。
  詳細は、「■ 設定手順」（→ P.39）を参照してください。

● 推奨理由

arp 監視モードは、STP を導入した環境では、mii 監視よりも監視できる範囲が狭くなり、適用できる形態も制限されるため、より制限のない mii 監視モードを推奨します。

■ 設定手順

ここでは、複数デバイスの場合の設定方法を記述していますが、bonding デバイスが 1 個だけの場合に適用しても問題はありません。

1 /etc/modprobe.conf ファイルを修正します。

以下の 1 行を追加します。

```
・
・
alias bond0 bonding
```

2 /etc/sysconfig/network-scripts/ifcfg-bond0 ファイルを追加します。

ファイルの内容は、以下のように設定します。

```
DEVICE=bond0
BOOTPROTO=static
IPADDR=xxx.yyy.zzz.www
NETMASK=255.255.255.0
NETWORK=xxx.yyy.zzz.0
BROADCAST=xxx.yyy.zzz.255
ONBOOT=yes
BONDING_OPTS="mode=1 primary=eth0 miimon=100"
```

bonding デバイス（master デバイス）の設定ファイルは、通常の NIC（ifcfg-ethX）の設定に加え、BONDING_OPTS 文で bonding オプションを記述します。
BONDING_OPTS 文に記述可能なオペランドは、表 3 - 2 のとおりです。

表 3 - 2：記述可能なオペランド

オペランド	説明
mode	bonding ドライバのモードを指定します。二重化機能を使用する場合は「1」固定です。
miimon	リンクの状態確認の間隔をミリ秒単位で指定します。省略すると監視が行われなくなるため、省略しないでください。特に要件がなければ「100」としてください。
primary	bonding デバイスを構成するスレーブインターフェースのうちの 1 つを優先して使用したい場合に、そのデバイス名を指定します。省略しても構いません。
updelay	リンクを有効にするまでの待ち時間を、miimon パラメーターの倍数（ミリ秒）で指定します。 スイッチが障害から回復したとき、リンクが不安定な状態にあると、NIC の切替え通知がスイッチに到達せず、結果として NIC の切替えに時間がかかる場合があります。この状態のとき、RHEL システムから NIC の切替え通知を送信しないようにすることで、短時間で NIC を切り替えることができます。 リンクが不安定になる場合以外、本設定は不要です。

3　インターフェース（eth2）を slave デバイスとする場合、/etc/sysconfig/network-scripts/ifcfg-eth2 ファイルを修正します。

以下のように修正します。

```
DEVICE=eth2
MASTER=bond0
SLAVE=yes
HWADDR=aa:bb:cc:dd:ee:ff
ONBOOT=yes
BOOTPROTO=static
TYPE=Ethernet
```

参考

- **追加した MASTER および SLAVE 行以外の記述内容は設定例です。通常、MASTER および SLAVE 行の追加以外にファイルを変更する必要はありません。**
- **VLAN と組み合わせる場合には、ifcfg-bond0 には IP アドレスを設定せず、ifcfg-bond0.X の VLAN 用の設定ファイルを追加してください。**

3.3.3　ロードバランシング機能

複数の NIC をグループ化することで、1 つの NIC のように見せかけて、帯域を拡張することで、ネットワークの負荷分散を実現する方法です。

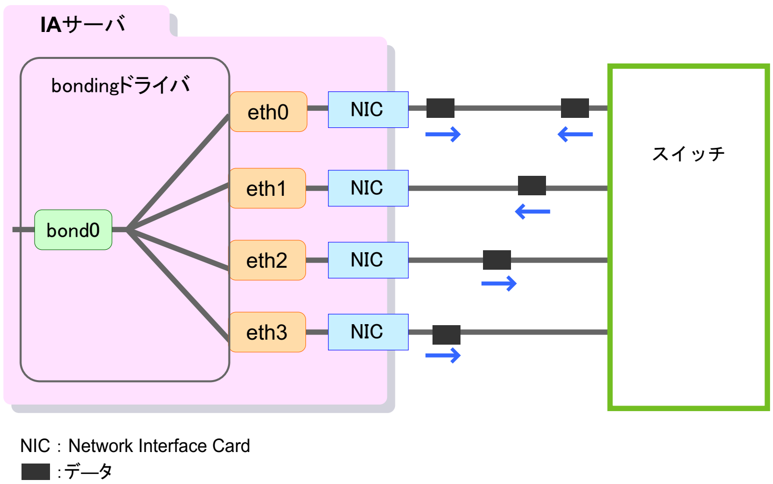


図 3 - 8：ロードバランシング機能の概要

ここでは、bonding ドライバを使用したロードバランシングを構築する場合の推奨モード、設定手順を説明します。

■ 推奨する設定

- 動作モード
  増設 LAN を必要な枚数（達成したい能力による）導入し、外部スイッチとの間を mode4（IEEE802.3ad）で接続することを推奨します。
- オプションの設定
  mii 監視モードを設定してください。
  無応答による異常検出の間隔のデフォルト値は、30 秒です。間隔を短くしたい場合は、「lacp_rate=1」を指定すると、異常検出の間隔が 1 秒になります。

● 推奨理由

各モードを送信方式、受信方式、制限で分類すると以下のようになります。

- 送信方式（送信データをどのリンクに割り当てるかの方式）

ラウンドロビン	mode0
相手単位に固定	mode2
複数の情報をもとに固定	mode4
動的に負荷分散	mode5、mode6

  相手単位に固定の方式は、1 つの通信相手のみを見ると、1 リンク以上に帯域を増やせないことを意味します （mode4、mode5、mode6 もこの点は同様です)。
- 受信方式（受信データをどのリンクに割り当てるか）

スイッチの動作による	mode0、mode2
相手単位に固定	mode6
複数の情報をもとに固定	mode4
全体で 1 リンクに固定	mode5

  mode0、mode2 の場合、動作の選択はスイッチに任されます。スイッチに設定がない場合、送信と受信で動作方式が非対称になる可能性もあります。ほとんどの場合、スイッチは、mode2 相当の動作を行います。
- 形態の制限

全リンクを 1 台のスイッチに収容する必要あり	mode0、mode2、mode4
複数のスイッチをまたいで動作可能	mode5、mode6

  mode0、mode2 の場合、動作の選択はスイッチに任されます。スイッチに設定がない場合、送信と受信で動作方式が非対称になる可能性もあります。ほとんどの場合、スイッチは、mode2 相当の動作を行います。

送信、受信方式とも mode4 は、mode0、mode2、mode5、mode6 を事実上包含しており、方式的にも最新です。したがって mode4 を推奨します。
mode5、mode6 はスイッチのサポートを必要とせず、かつ、リンクを複数のスイッチに分散させたい（二重化を同時に実現したい）場合は選択肢となりえます。この場合は、機能的に上位の mode6 を使用してください。

■ 設定手順

ここでは、複数デバイスの場合の設定方法を記述していますが、bonding デバイスが 1 個だけの場合に適用しても問題はありません。

1 /etc/modprobe.conf ファイルを修正します。

以下の 1 行を追加します。

```
•
•
alias bond0 bonding
```

2 /etc/sysconfig/network-scripts/ifcfg-bond0 ファイルを追加します。

ファイルの内容は、以下のように設定します。

```
DEVICE=bond0
BOOTPROTO=static
IPADDR=xxx.yyy.zzz.www
NETMASK=255.255.255.0
NETWORK=xxx.yyy.zzz.0
BROADCAST=xxx.yyy.zzz.255
ONBOOT=yes
BONDING_OPTS="mode=4 miimon=100"
```

bonding デバイス（master デバイス）の設定ファイルは、通常の NIC（ifcfg-ethX）の設定と変わりません。ただし、物理 MAC アドレスを持たないので HWADDR を記載する必要はありません。

参 考

- **上記のファイル内容は設定例です。ネットワーク構成などに依存して異なります。必要に応じて、RedHat 社のマニュアルを参照して変更してください。**

3 インターフェース（eth2）を slave デバイスとする場合、/etc/sysconfig/network-scripts/ifcfg-eth2 ファイルを修正します。

以下のように修正します。

```
DEVICE=eth2
MASTER=bond0
SLAVE=yes
HWADDR=aa:bb:cc:dd:ee:ff
ONBOOT=yes
BOOTPROTO=static
TYPE=Ethernet
```

参 考

- **追加した MASTER および SLAVE 行以外の記述内容は設定例です。通常、MASTER および SLAVE 行の追加以外にファイルを変更する必要はありません。**
- **VLAN と組み合わせる場合には、ifcfg-bond0 には IP アドレスを設定せず、ifcfg-bond0.X の VLAN 用の設定ファイルを追加してください。**

3.3.4 bonding ドライバと VLAN の組み合わせ

■ 利用目的と推奨する組み合わせ

bonding ドライバを使用したネットワークの設計は、動作モード／監視モード／ VLAN 接続の組み合わせで、様々なパターンが考えられます。
ここでは、利用目的別に推奨される組み合わせパターンを表 3 - 3 に示します。

表 3 - 3：利用目的に対する推奨パターン

利用目的	動作モード	監視モード	VLAN 接続
二重化機能	mode1	mii 監視	可能
ロードバランシング機能 ［注 1］	mode4 ［注 2］		

［注 1］：PRIMEQUEST 1000 シリーズのみサポートされます。
［注 2］：動作モードを mode4 で設計する場合は、接続するネットワークスイッチも IEEE802.3ad 仕様に準拠しているものを用意してください。

■ VLAN ネットワークとの組み合わせ時の注意

bonding ドライバを使用した二重化機能／ロードバランシング機能と、VLAN を組み合わせる場合の注意事項について説明します。

● bonding ドライバが結合するインターフェースについて

bonding ドライバが結合するインターフェースは、任意の NIC を選べますが、これらは、物理デバイスでなければなりません。例えば、bonding ドライバと tagVLAN デバイスが共存することは可能ですが、この場合、必ずスレーブインターフェースとして、ethX、ethY を選択し、それらを束ねた bondZ 上に bondZ.10 などを作成するようにしてください。スレーブインターフェースとして ethX.10、ethY.10 などの論理デバイス（VLAN デバイス）を選択する構成はサポートされません。

● mii 監視モードと VLAN

- mii 監視モードで VLAN を設定しない場合
  特に注意事項はありません。
  portVLAN を使用する場合は、サーバには VLAN の設定は不要ですが、リンクを収容するスイッチのポートが portVLAN として構成される場合は、ネットワーク全体では、VLAN 設定ありとして設計を行ってください。
- mii 監視モードで VLAN を設定する場合
  mii 監視モードで VLAN を設定する場合には、VLAN を設定したスイッチを冗長化し、STP の拡張である MSTP プロトコルをサポートするスイッチを用意して、スイッチの MSTP プロトコル機能を有効にしてください。

3.3.5 コマンドによる bonding デバイスの操作

bonding デバイスの状態は、以下の 3 種類のコマンドで変更できます。
本項では、bonding デバイス本体と、構成する個々のネットワークインターフェースを対比して記述する場合、master デバイス、slave デバイスという表現をします。この場合、master デバイスは、bonding デバイスを意味します。

- ifconfig
- ifup ／ ifdown
- ifenslave

ifconfig、ifup、ifdown には、操作の対象とするデバイスを 1 つ指定します。これは、master デバイス、slave デバイスのいずれであってもかまいません。ifenslave は、bonding デバイスの構成、状態の変更を行うコマンドであり、master デバイス、slave デバイスの両方を同時に指定します（詳細は指定するオプションにより異なります）。
それぞれのコマンドは、以下のように動作します。

● ifconfig

通常のデバイスと同様に、master デバイス、slave デバイスともに、活性／非活性の設定、IP アドレスの付与などを行います。ただし、bonding デバイスの構成の変更はできません。後述の ifup ／ ifdown、ifenslave を使用してください。ifconfig コマンドの詳細および使用例は、man マニュアルを参照してください。

● ifup ／ ifdown

bonding デバイス全体（master デバイス）、個々の slave デバイスの活性／非活性を行います。
操作対象を master デバイスか、slave デバイスとした場合で動作が異なります。
また、ifconfig と異なり、動作させるためには、対応する設定ファイル（ifcfg-bondX、ifcfg-ethX）が必要です。

master デバイスを対象にした場合

ifup を実行すると、ifcfg-ethX ファイル群を検索して、対象の master デバイスに属するすべての slave デバイスを組み込んで活性化します。さらに、master デバイスを活性化し、必要に応じて IP アドレスの割り当てなどを ifcfg-bondX ファイルの記述に従って行います。
ifdown を実行すると、master デバイスに属するすべての slave デバイスを取り外して、非活性化します。さらに master デバイスも非活性化します。

slave デバイスを対象にした場合

ifup を実行すると、対象となる ifcfg-ethX ファイルから master デバイスを見つけ、slave デバイスとして組み込んで活性化します。
ifdown を実行すると、slave デバイスを master デバイスから取り外して、非活性化します。
ifup ／ ifdown コマンドの詳細および使用例は、RedHat 社の公開するマニュアルを参照してください。

● ifenslave

master デバイスへの slave デバイスの組み込み／取り外し、アクティブな slave デバイスの変更などを行います。主に、一時的に構成を変更するために使用します。詳細は、man マニュアルを参照してください。
以下に使用例を示します。

master デバイス「bond0」に slave デバイス「eth2」を追加する例

```
# ifenslave bond0 eth2
```

master デバイス「bond0」から slave デバイス「eth2」を取り外す例

```
# ifenslave -d bond0 eth2
```

master デバイス「bond0」でアクティブ状態となっているスレーブインターフェースを「eth1」に切り替える例

```
# ifenslave -c bond0 eth1
```

ifenslave は、master デバイスおよび slave デバイスを直接指定できるので、ifcfg-ethX ファイルの内容に関わらず、任意の構成を作成できます。
また、ifenslave の結果は、システムを再起動すると無効となるため、恒久的に変更を行いたい場合は、ifcfg-bondX および ifcfg-ethX を追加、変更する必要があります。

第 4 章

システム設計：デバイス名ずれ対策設計

この章では、デバイス名ずれの要因およびデバイス名ずれに対応するために必要な設計について説明します。

4.1 ハードウェア構成変更を考慮したデバイス名の運用設計

ハードウェア構成変更を考慮したデバイス名の運用設計について説明します。

● デバイス名とは？

Linux では、ディスク装置やテープ装置などの周辺機器に対して入出力を行うために、デバイスファイルと呼ばれるものを利用します。このファイル名のことをデバイス名といい、これを指定して入出力処理を行うことにより、対応する周辺機器に実際の入出力が行われます。
デバイスファイルはシステム起動時に /dev ディレクトリ配下に作成され、システムから認識された順番に名前が付けられます。
末尾には、ユニットと呼ばれるアルファベット（a,b,c,…）または数字（0,1,2,…）が割り当てられます。

デバイス名の例

表 4 - 1：デバイス名の例

デバイス名	対応する周辺機器	システムから認識された順番
/dev/sda	SCSI デバイス	1 番目に認識
/dev/sdb	SCSI デバイス	2 番目に認識
/dev/sdc	SCSI デバイス	3 番目に認識

4.1.1 デバイスの認識方法とデバイス名の決まり方

Linux では、ハードディスクドライブなどの各種デバイスに対し、システム起動時にシステムが認識した順序でデバイス名を割り当てます。デバイスの認識順序は、以下のようになります。

1 ドライバの組み込み順序

ドライバは以下の順序で組み込まれます。

1. **/etc/fstab ファイルでルートファイルシステムに指定されているデバイスに対応するドライバ**
2. **/etc/fstab ファイルで swap 領域に指定されているデバイスに対応するドライバ**
3. **上記以外のドライバは /etc/modprobe.conf ファイルの alias 行の指定順**

2　同一ドライバを使用する PCI デバイスの検出順序

同一ドライバを利用する PCI デバイスの検出順序は、PCI デバイスの構成方法に依存します。そのため、単純にバス番号順にはならないことがあります。

4.1.2　デバイス名に関するトラブル

デバイス名は、起動時のデバイスの認識順序が前回起動したときと異なる場合に変わることがあります。これを「デバイス名ずれ」といいます。

例えば、ハードウェア故障やデバイスの交換後、再起動した時点でハードディスクドライブやコントローラーなどの一部がシステムから認識できなくなった場合に、デバイス名が変更されることがあります。これにより、再起動前に使用していたデバイス名が、別のデバイスになっていることがあり、データ破壊などのトラブルにつながります。

「デバイス名ずれ」が発生することにより、システムが起動できないトラブルや、データ参照ができないトラブルが想定されます。
デバイス名ずれが発生するデバイス、および発生要因は以下のとおりです。

■ デバイス名ずれの発生するデバイス

デバイス名ずれは、以下の 2 種類のデバイスで発生します。

- ディスク系デバイスで発生するもの
  ディスク（ハードディスクドライブやコントローラーなど）
- ネットワーク系デバイスで発生するもの
  NIC（ネットワークインターフェースカード）

■ デバイス名ずれの要因

デバイス名ずれは、例えば以下に示す要因により発生します。

- 「4.1.3 デバイス名ずれの要因：ソフトウェア変更」（→ P.48）
- 「4.1.4 デバイス名ずれの要因：ハードウェア構成変更」（→ P.51）
- 「4.1.5 デバイス名ずれの要因：ハードウェア故障」（→ P.55）

また、システムの再起動でもデバイス名ずれは発生します。

4.1.3　デバイス名ずれの要因：ソフトウェア変更

ソフトウェア変更により発生するデバイス名ずれの原因は、以下のものがあります。

- 「■ ドライバの定義変更」（→ P.49）
- 「■ カーネルアップデート時のドライバ適用状況」（→ P.49）

これらは、デバイスの認識順序の「ドライバの組み込み順序」に関係しています。
ドライバの組み込み順についての詳細は「4.1.1 デバイスの認識方法とデバイス名の決まり方」（→ P.47）を参照してください。

■ ドライバの定義変更

/etc/modprobe.conf ファイルの alias 行の修正・追加・削除によりドライバの組み込み順が変更されると、デバイス名ずれが発生する可能性があります。

例：ドライバ定義変更によるデバイス名ずれ

/etc/modprobe.conf ファイルを編集して、システムが起動するときにドライバが組み込まれる順序を入れ替えた場合を例とします。
この例では、内蔵ディスクがシステムディスク（システム制御に必要なファイルが入っているディスク）のため、デバイス名ずれの発生により、システムが起動できなくなります。

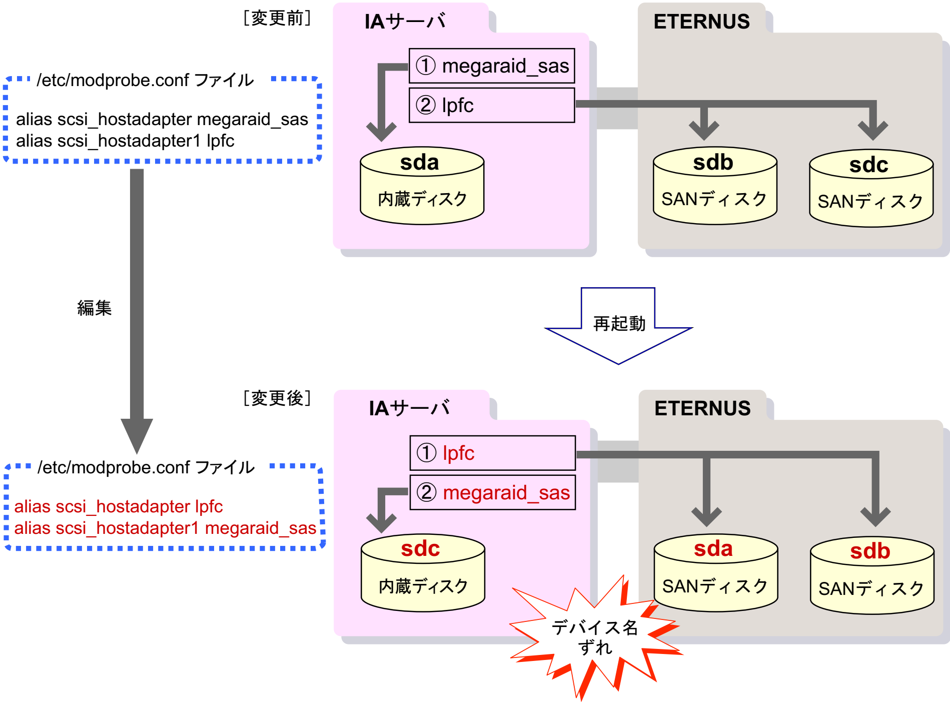


図 4 - 1：ドライバの定義変更によるデバイス名ずれの例

■ カーネルアップデート時のドライバ適用状況

RHEL で標準提供されないドライバ（富士通提供ドライバなど）を用いている場合、カーネルに適合した版数のドライバを使用する必要があります。カーネルアップデートの際は、RHEL で標準提供されないドライバも、アップデートしたカーネルに対応する版数のドライバ適用が必要となる場合があります。カーネルアップデートにより RHEL で標準提供されないドライバの版数がカーネルの版数に適合しなくなった場合、そのままの状態でシステムが起動されると、当該ドライバの組み込みに失敗します。この場合、対応するデバイスが認識できなくなり、デバイス名ずれが発生します。

例：カーネルアップデート時の富士通提供ドライバ未更新によるデバイス名ずれ

カーネルアップデート後に、富士通提供ドライバ（以下の例では富士通提供の mptbase ドライバ）を更新しない状態のまま、システムを起動した場合を例とします。
この場合、アップデートしたカーネルでは富士通提供ドライバがロードされず、ハードディスクキャビネット内のディスクが認識できないため、SAN Disk（ETERNUS）のデバイス名がずれます。

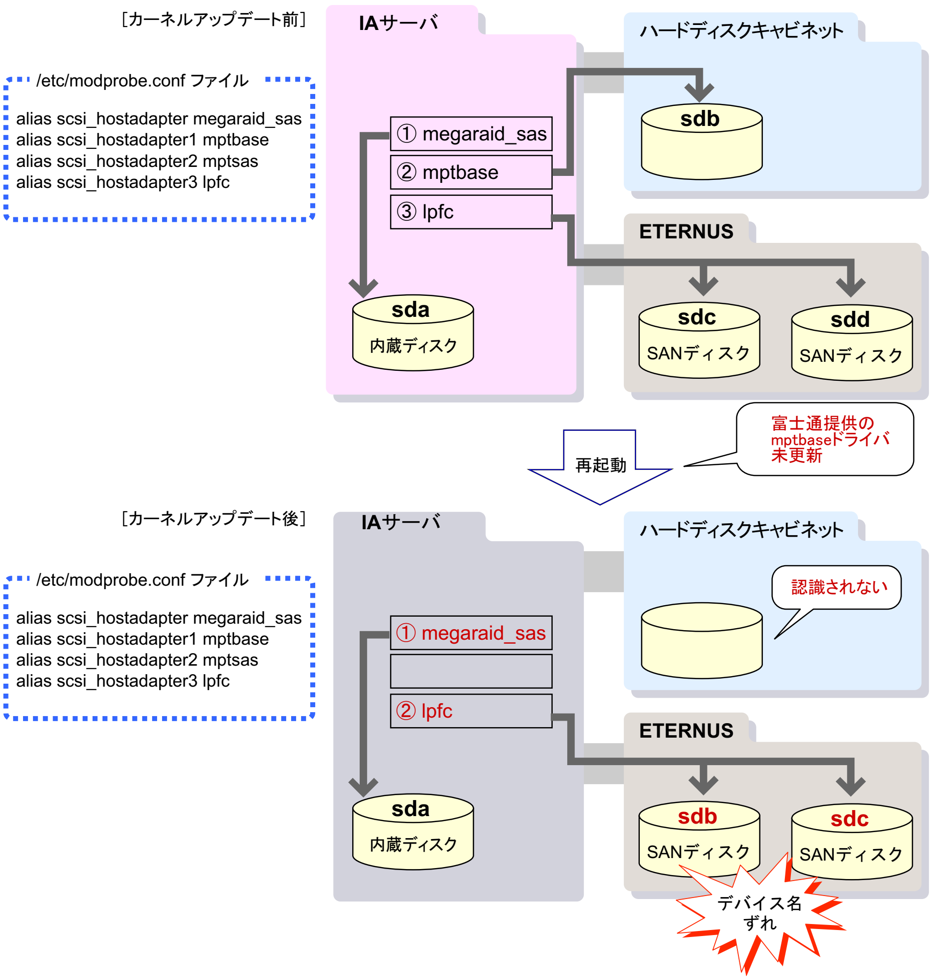


図 4 - 2：カーネルアップデートによるデバイス名ずれの例

4.1.4 デバイス名ずれの要因：ハードウェア構成変更

ハードウェア構成変更により発生するデバイス名ずれの原因は、以下のものがあります。

- 「■ ディスク（リムーバブルディスク）の追加・削除」（→ P.51）
- 「■ PCI スロットへのカード追加・削除」（→ P.53）
- 「■ NIC の増設・削除・交換、または LAN ポートを持つシステムボードの交換」（→ P.54）

これらは、デバイスの認識順序の「PCI デバイスの検出順序」に関係しています。
PCI デバイスの検出順序についての詳細は「4.1.1 デバイスの認識方法とデバイス名の決まり方」（→ P.47）を参照してください。

■ ディスク（リムーバブルディスク）の追加・削除

接続形態（サーバ本体内蔵／外付けストレージ装置）や装置種別（ハードディスク装置など）にかかわらず発生する可能性があります。保守作業でディスクの追加や削除などの構成変更を行った場合、再起動後に既存のディスクのデバイス名がずれることがあります。

例：ディスクの追加によるデバイス名ずれ

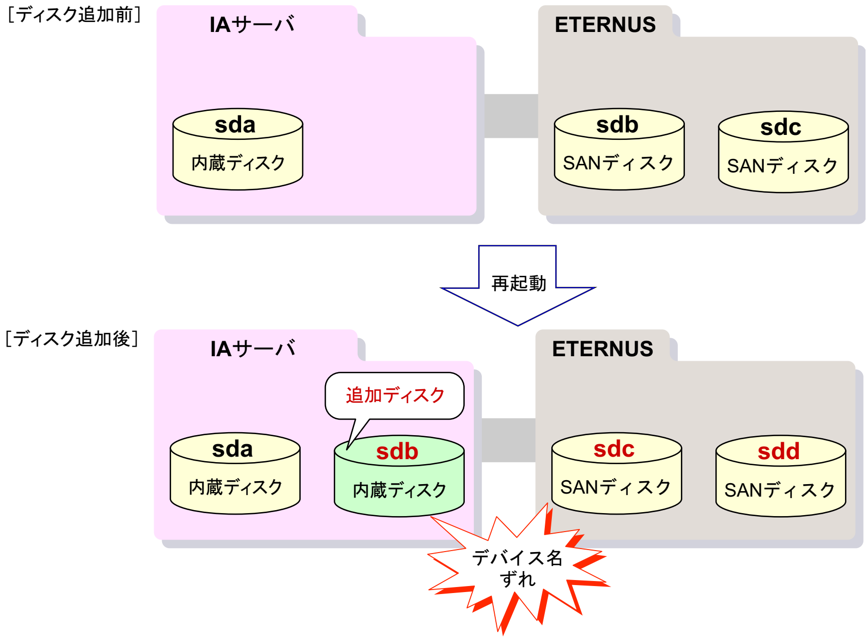


図 4 - 3：ディスクの追加によるデバイス名ずれの例

例：デバイス（LU：LogicalUnit）の追加によるデバイス名ずれ

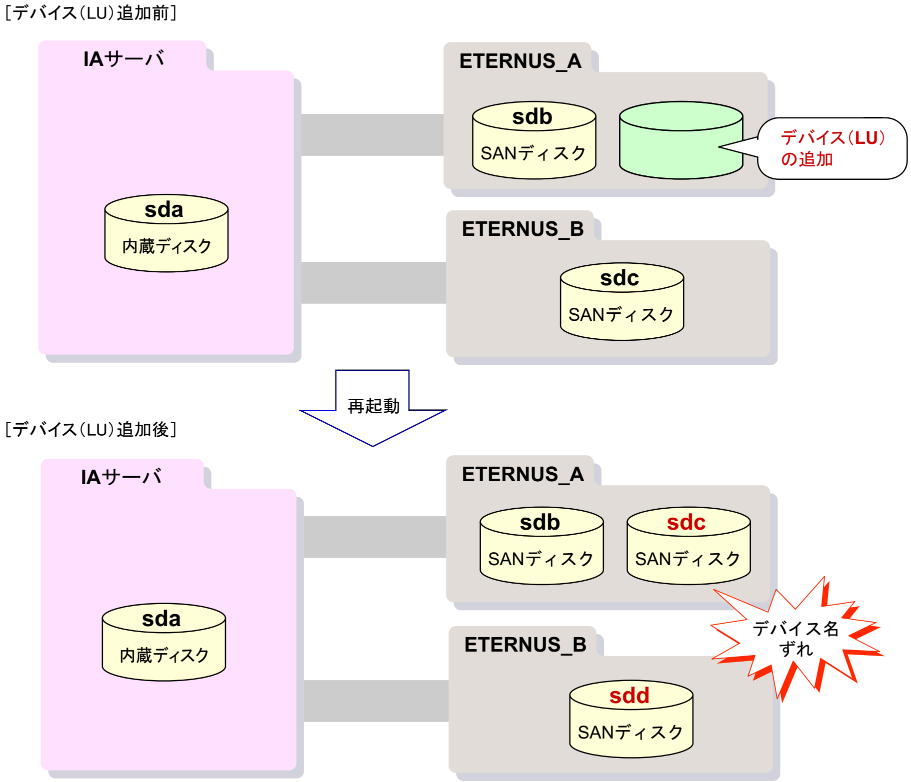


図 4 - 4：デバイス（LU：LogicalUnit）の追加によるデバイス名ずれ

注 意

ハードウェア構成変更の注意事項

- ディスクの追加や削除などの構成変更を行ったあと、最初のシステム起動時のファイルシステムチェックでディスクの構成情報が不整合となり、システムの起動に失敗する場合があります。
  これは、fsck コマンドが参照するデバイス構成情報のキャッシュファイルがファイルシステムチェック時には更新されていないために、fsck コマンドはディスクの追加や削除を行う前の古いデバイス構成情報をもとにファイルシステムをチェックするためです。
  ディスクの追加や削除などの構成変更を行う場合は、ディスクの追加や削除を行う前のシステム停止直前に、キャッシュファイルを削除してください。キャッシュファイルのパスは以下のとおりです。

```
/etc/blkid/blkid.tab
```

■ PCI スロットへのカード追加・削除

PCI デバイスの検出順序は、単純にバス番号順にならないことがあります。詳しくは「4.1.1 デバイスの認識方法とデバイス名の決まり方」(→ P.47)を参照してください。そのため、対象とする機種のハードウェアマニュアルにおいて、PCI スロットへのオプションカードの追加、取り外し手順箇所には、オプションカードの PCI スロットへの搭載順が提示されている場合があります。この搭載順に従わずにオプションカードを PCI スロットへ搭載すると、デバイス名ずれが発生することがあります。

例：PCI スロットへのカード追加によるデバイス名ずれ

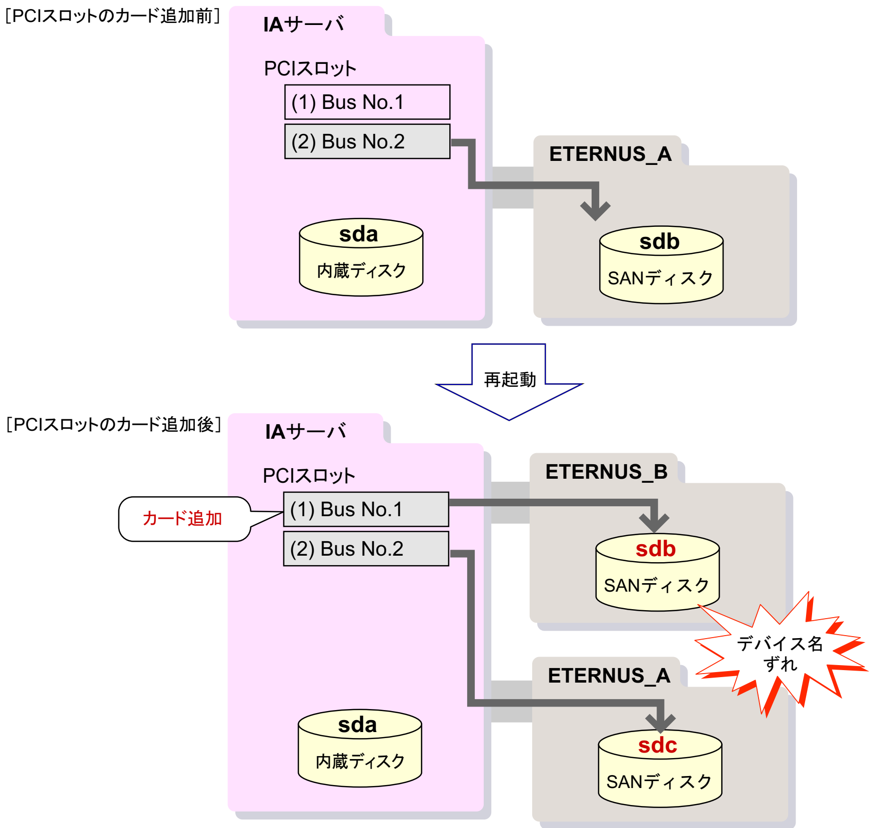


図 4 - 5：PCI スロットへのカード追加によるデバイス名ずれの例

ハードウェアマニュアルに記載されている PCI スロットへのオプションカードの搭載順が (1) → (2) の場合、(2) のスロットから先に搭載して運用を開始し、その後 (1) の PCI スロットにカードが追加されると、デバイスの認識順序が変わり、「ETERNUS_A」のデバイス名がずれることがあります。

カードを取り外した場合は、追加と逆の現象となりデバイス名が変更されます。

■ NIC の増設・削除・交換、または LAN ポートを持つシステムボードの交換

NIC またはシステムボードの機器構成が変更された場合、ネットワークデバイス名に影響を及ぼすことがあります。

● NIC 増設時のネットワークデバイス名ずれ

NIC を増設した場合、ネットワークデバイス名は、増設したカードが「eth0」、「eth1」、オンボード LAN が「eth2」、「eth3」となることがあります。

例：NIC 増設時のデバイス名ずれ

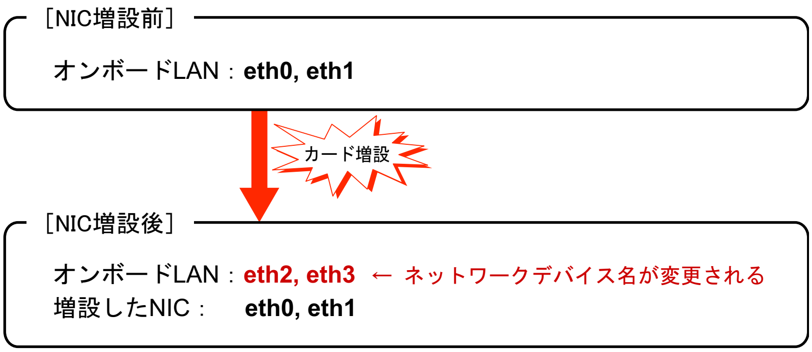


図 4 - 6：NIC 増設時のデバイス名ずれ

NIC を取り外した場合も、図 4 - 6 と逆となり、ネットワークデバイス名が変更されることがあります。

● 複数 LAN ポートを有する NIC またはシステムボードの交換によるデバイス名ずれ

複数の LAN ポートを持つ NIC やシステムボードを交換した場合、交換後の再起動にて、LAN ポートとネットワークインターフェース (ethX,ethY, ...) との対応が、交換前の状態と異なってしまう場合があります。これは、LAN ポートの MAC アドレスに一致する ifcfg-ethX ファイルが存在しない場合、システムはその LAN ポートに対して、順不同でネットワークデバイス名を割り当てるためです。

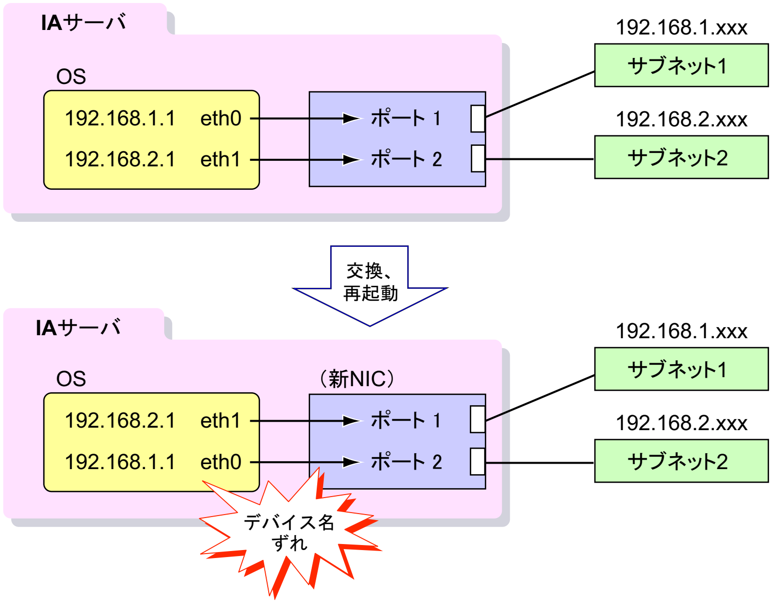


図 4 - 7：NIC またはシステムボードの交換によるデバイス名ずれ

4.1.5 デバイス名ずれの要因：ハードウェア故障

ハードウェアが故障した場合も、ハードウェアを削除した場合と同様にシステムの再起動時にデバイス名ずれが発生します。
ハードウェア故障により発生するデバイス名ずれの原因は、以下のものがあります。

- 「■ ディスク故障」（→ P.55）
- 「■ カード故障」（→ P.56）
- 「■ ストレージ装置故障・電源断・ケーブル故障」（→ P.56）
- 「■ NIC 故障」（→ P.58）

■ ディスク故障

システムボリューム (/dev/sda) が故障し、デバイスがシステムから認識できなくなった場合は、システムが起動しないため、デバイス名ずれの予防保守対策はできません。
システムに 3 台のハードディスクドライブが接続されている場合、それぞれデバイスの検出順にデバイス名として、sda、sdb、sdc が割り当てられます。
この状態で sdb が故障し、次の起動時に元々 sdb であったデバイスがシステムから認識できなくなった場合、元々 sdc であったデバイスには、デバイス名として sdb が割り当てられます。

例：ディスク故障によるデバイス名ずれ

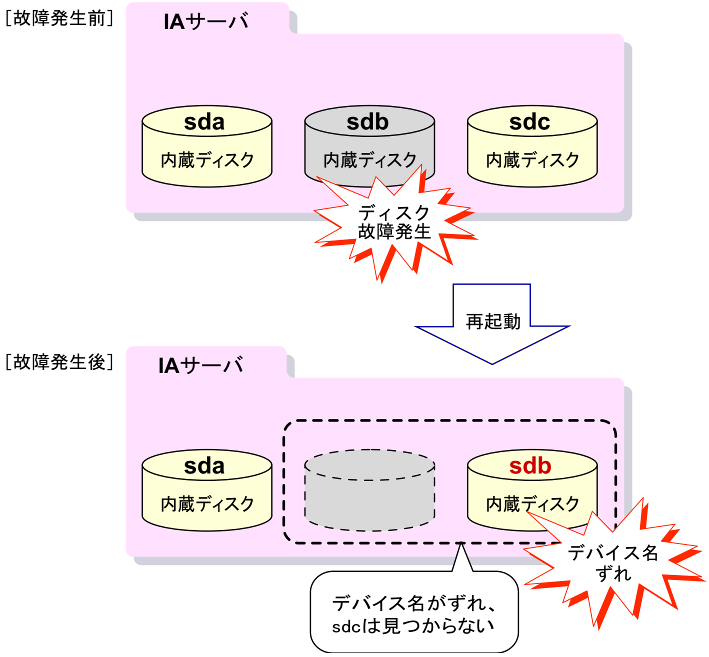


図 4 - 8：ディスク故障によるデバイス名ずれ

■ カード故障

カードが故障し、システムから認識できなくなった場合には、カードを削除した場合と同じ現象となります。

■ ストレージ装置故障・電源断・ケーブル故障

ストレージ装置内のデバイスがシステムから認識できなくなった場合に発生します。
なお、デバイス名ずれは、ハードウェア故障要因の他にも発生する可能性があります。筐体の電源切断や接続ケーブルの断線によっても、デバイス名ずれが発生します。

例：ストレージ装置故障によるデバイス名ずれ

［故障発生前］

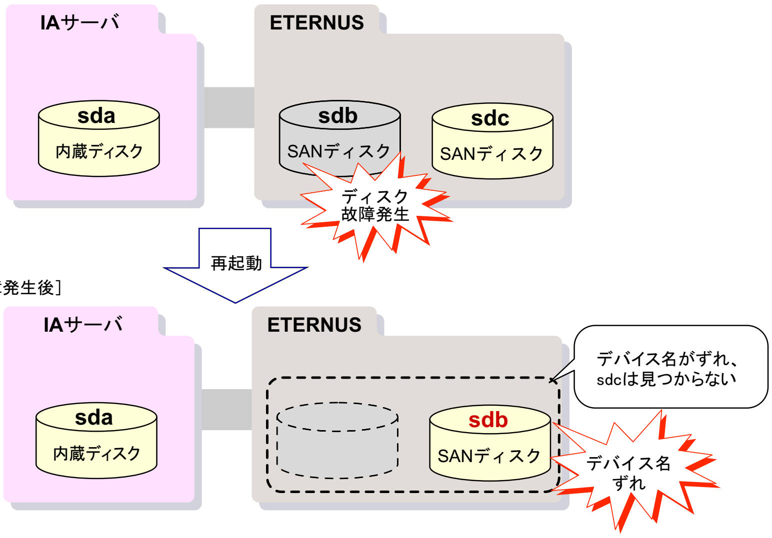


図 4 - 9：ストレージ装置故障によるデバイス名ずれ

例：電源断／ケーブル切断によるデバイス名ずれ

［ケーブル切断前］

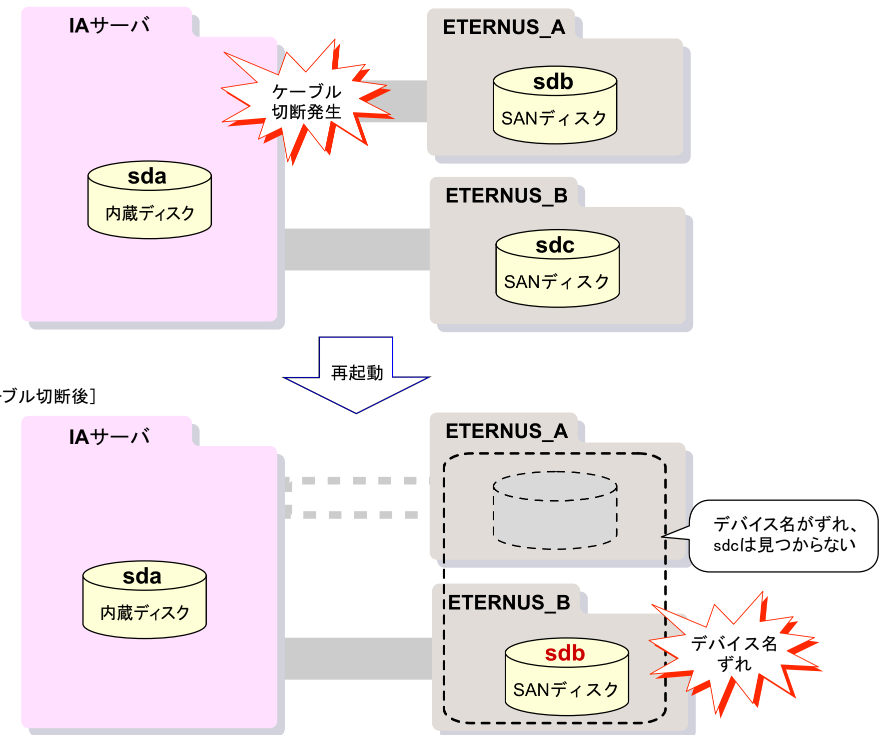


図 4 - 10：電源断／ケーブル切断によるデバイス名ずれ

■ NIC 故障

NIC が故障し、システムから認識できなくなった場合には、PCI カードかオンボード LAN かを問わず、NIC を削除した場合と同じ現象により、デバイス名ずれが発生します。

例：NIC 故障時のデバイス名ずれ

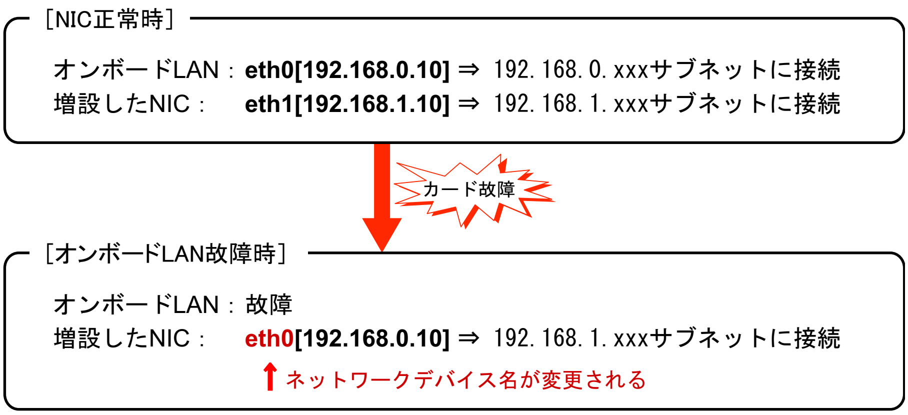


図 4 - 11：NIC 故障時のデバイス名ずれ

4.1.6 デバイス名が変わることによる影響

デバイス名ずれは、すべてのディスクやネットワークデバイスで発生する可能性があります。ただし、システムの機器構成、故障発生箇所の違いにより、デバイス名ずれ自体やデータ破壊を引き起こす度合いが異なります。そのため、対象となるシステムにどの程度のリスクがあるのかを把握することが重要です。

万が一システムからデバイスが認識できなくなった場合の影響は、以下のとおりです。

■ 運用停止

システムディスクがディスク故障となった場合は、システムが起動できなくなります。

■ ハードディスクのデータ破壊

データ用のディスクに故障が発生した場合は、以下のようにデバイス名ずれによるデータ破壊を引き起こす可能性があります。

- 「● ファイルシステムとして利用するディスクが存在する場合」（→ P.59）
- 「● raw デバイスとして利用するディスクの場合」（→ P.60）

● ファイルシステムとして利用するディスクが存在する場合

デバイス名ずれにより、意図しないパーティションをマウントする可能性が考えられます。そのため、最悪の場合はデータ破壊が発生します。
ただし、ラベル指定でマウントする場合は、誤って他のパーティションをマウントすることはありません。ラベル指定については「4.2.2 ディスク系デバイスの対処方法（ラベル名による運用）」（→ P.64）を参照してください。

例：デバイス名ずれ発生時の影響（ファイルシステム）

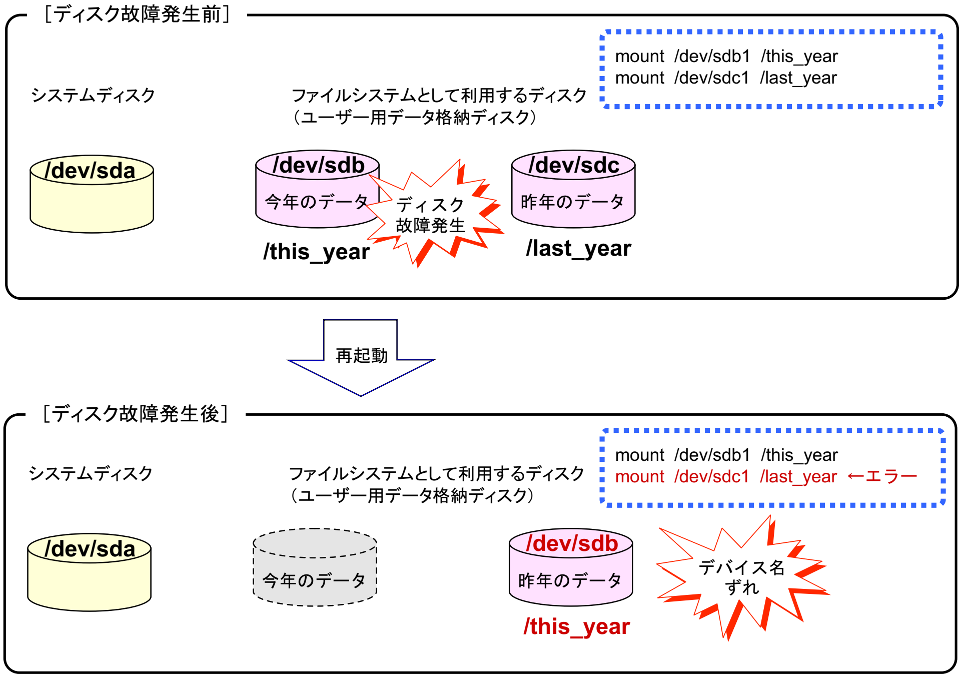


図 4 - 12：デバイス名ずれ発生時の影響（ファイルシステム）

今年のデータが格納されているディスクが故障したため、デバイス名ずれにより、昨年のデータを格納するディスクが /this_year でマウントされてしまいます。
今年のデータのディスクと昨年のデータのディスクのパーティション構成が同じ場合は、昨年のデータを今年のデータとして参照することで、データ破壊が発生する場合があります。

● **raw デバイスとして利用するディスクの場合**

デバイス名の変更により、別のディスクに bind される可能性があり、最悪の場合、データ破壊が発生します。

例：デバイス名ずれ発生時の影響（raw デバイス）

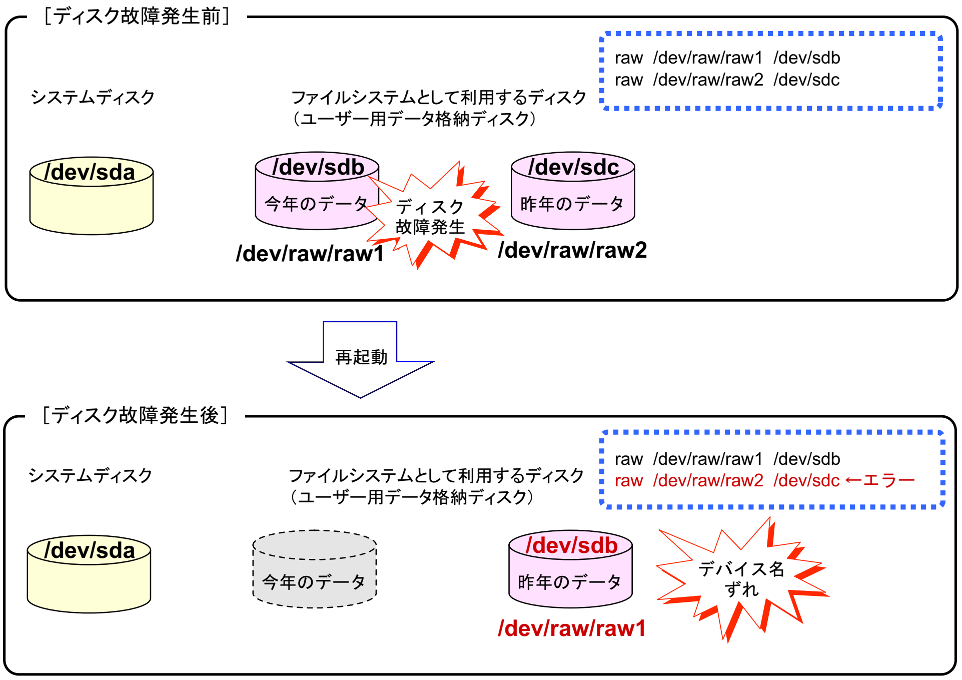


図 4 - 13：デバイス名ずれ発生時の影響（raw デバイス）

今年のデータが格納されているディスクが故障したため、デバイス名ずれにより、昨年のデータを格納するディスクに /dev/raw/raw1 が bind されます。昨年のデータを今年のデータとして参照することによりデータ破壊が発生します。

■ ネットワーク環境

Linux のデバイス認識方式に起因した不具合により、業務アプリケーションやミドルウェア、またはシステム（ネットワーク環境）に影響を与える場合があります。
ネットワーク系デバイスでは、NIC の増設や故障時に、ネットワークデバイス名（ethX）が変更されることにより、静的 IP アドレス管理を行っている環境では、異なる IP アドレスが物理的な LAN ポートに割り当てられるなどの不具合が発生し、通信ができない状況になる場合があります。

例：NIC 故障時のデバイス名ずれ

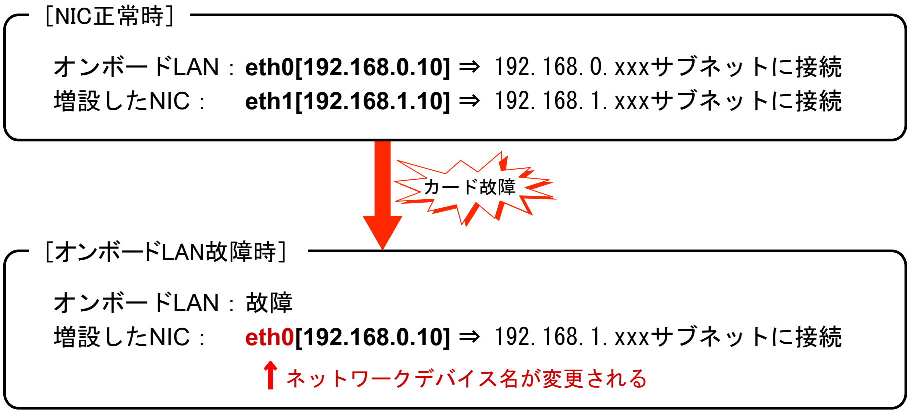


図 4 - 14：NIC 故障時のデバイス名ずれ

この不具合は、ハードウェアアドレスとネットワークデバイス名を固定にすることで対処できます。対処方法については、「4.3.2 ネットワーク系デバイスの交換・増設時の注意」（→ P.78）を参照してください。

■ 影響を受けるコマンド、設定ファイル

引数にデバイス名の指定が必要なコマンドや、デバイス名が記述される設定ファイルは、デバイス名ずれが発生すると正常に動作しない場合があります。
影響を受ける主要なコマンドおよびファイルを示します。

表 4 - 2：デバイス名ずれで影響を受ける主要なコマンド／設定ファイル

操作の目的	コマンド	設定ファイル
パーティションの設定	parted, fdisk	/etc/fstab
ブートローダの設定	grub-install	/boot/grub/grub.conf
ファイルシステムの操作	mkfs, mkfs.ext2, mkfs.ext3, e2label, fsck, e2fsck, tune2fs, mount, umount	/etc/fstab, /etc/auto.master
バックアップ・リストア	dump, restore	
ダンプ設定		/etc/kdump.conf
Logical Volume Manager（LVM）の操作	pvcreate, vgcreate	/etc/lvm/cache, /etc/lvm/lvm.conf
raw デバイスの操作	raw	/etc/udev/rules.d/60-raw.rules
swap ファイルの操作	mkswap, swapon, swapoff	/etc/fstab
ネットワークインターフェースの操作	ifconfig, ifup, ifdown	/etc/sysconfig/network-scripts/ifcfg-ethX

4.2 設計・導入時の考慮

ハードウェアが故障した場合に発生するデバイス名ずれに備えるために設計・導入時に考慮すべき点について説明します。

4.2.1 ディスク系デバイスの対処方法

■ ハードウェアによる防止策

ハードウェアによる防止策として、RAID 機能によるハードディスクの冗長化があります。ハードディスクを冗長構成にすると、単一のハードディスク故障が発生してもハードディスクの内容をシステムから参照することができます。
システムディスクは、ハードウェア RAID（RAID1 など）を使用するようにしてください。ただし、ハードウェア RAID によるディスク冗長化を実施しても、ディスク故障によるデータ損失は防げますが、デバイス名ずれの発生を解消するものではありません。

● ハードウェア故障時のシステム起動停止による対処例

データ破壊を防止する方法として、ハードディスク系デバイスが故障している場合に、システム起動を停止させる仕組みを導入することが有効です。
この仕組みは、システム起動時に動作する、/etc/rc.d/rc.local ファイルなどにより行うことができます。
デバイスの故障を判断するシェルスクリプトを用意し、/etc/rc.d/rc.local ファイルから呼び出す設定とします。このシェルスクリプトは、正常にシステムが起動した状態のデバイス構成情報（/proc/partitions ファイル）を保存しておき、次回システム起動時のデバイス構成情報に変更がないことを判断します。デバイス構成情報が変更されている場合は、オペレーター介入を求めるメッセージを出力し、システムを停止状態とするかを問い合わせます。

デバイスの故障を判断するシェルスクリプトの例は、以下のとおりです。

```
#!/bin/sh
# diskcomp_chk.sh
# This is a script which checks whether the disk composition has changed.
# It notifies the operator when there is a change in the disk composition.
PROC=/proc/partitions
COMP_INFO=/etc/disk_composition
TEMP=/tmp/disk_composition
inquiry() {
        yn_result=""
        while [ -z "$yn_result" ]; do
                read yn_reply
                case "$yn_reply" in
                        [Yy][Ee][Ss]) # yes
                                yn_result=0;;
                        [Nn][Oo]) # No
                                yn_result=1;;
                        *)
                        echo "Please enter yes or no.";;
                esac
         done
         return $yn_result
}
if [ -f $COMP_INFO ]
then
         cat $PROC | sed -e 's/ \+/ /g' | cut -d' ' -f2-5 > $TEMP
         RESULT=`diff -N $COMP_INFO $TEMP | head -n 1`
         if [ -z $RESULT ]
         then
                 # There is no change.
                 exit
         else
                 echo "There is a change in the composition."
                 echo "Does the system down? (yes/no)"
                 if inquiry; then
                          echo " Please system down : /sbin/shutdown -h now"
                 else
                          cat $TEMP > $COMP_INFO
                 fi
         fi
else
         echo "create $COMP_INFO"
         cat $PROC | sed -e 's/ \+/ /g' | cut -d' ' -f2-5 > $COMP_INFO
fi
```

■ 運用上の防止策

システム構成によっては、一部のデバイスが参照できない状態（閉塞状態）でもシステムを停止せずに稼働を継続できる場合があります。
その際に、デバイス名ずれの発生によってデータ破壊が引き起こされないように、ソフトウェア（OS またはミドルウェア製品）の機能を用いた、以下の対策をとる必要があります。

- 「4.2.2 ディスク系デバイスの対処方法（ラベル名による運用）」（→ P.64）
- 「4.2.3 ディスク系デバイスの対処方法（udev 機能）」（→ P.67）
- 「4.2.4 ディスク系デバイスの対処方法（ミドルウェア機能）」（→ P.71）

4.2.2 ディスク系デバイスの対処方法（ラベル名による運用）

いくつかのファイルシステムや swap 領域にはラベル名（ファイルシステムや swap 領域が格納しているデータを独自に認識するのに使用できる文字列）を保存する機能があります。
ラベル名を使用して、ファイルシステムや swap 領域を操作する機能（以降、ラベル指定と表記します）は、ファイルシステムをマウントしたり swap 領域を有効にしたりする際に使用できるため、デバイス名を使用する必要がなくなります。
システム起動時にも、ラベル指定を使用できるため、万が一デバイス名ずれが発生したとしても、誤ったデバイスを使用することはありません。
ファイルシステム、または swap 領域として利用するパーティションには、ラベル名を設定するようにしてください。

■ ラベル名を設定する

システムの初期インストール時にディスクを初期化したシステムでは、デフォルトでパーティションにラベル名が付与されます。
手動でラベル名を設定する場合、システム全体で固有なラベル名を設定する必要があります。

ファイルシステム、swap 領域によって設定コマンドが異なります。該当する手順を参照してください。

- 「● ファイルシステムが ext2、ext3 の場合」（→ P.65）
- 「● swap 領域の場合」（→ P.65）

注 意

- **同一ラベル名のファイルシステムが複数あると、目的のファイルシステムにアクセスができないことがあります。このため、ファイルシステムの種類に合ったコマンドを実行してシステム全体で固有なラベル名を設定してください。**
- **ファイルシステムのラベル名を設定する場合は、ファイルシステムがアンマウントされた状態で行ってください。**

● ファイルシステムが ext2、ext3 の場合

以下の手順では、＜ device ＞を「/dev/sdb1」、＜ label ＞を「DATA_2010」にした場合を例に説明します。

1 **以下のコマンドを実行し、ラベル名を設定します。**

```
# /sbin/e2label /dev/sdb1 DATA_2010
```

2 **以下のコマンドを実行し、ラベル名を確認します。**

```
# /sbin/e2label /dev/sdb1
```

手順 1 で設定したラベル名「DATA_2010」が、表示されます。

3 /etc/fstab **ファイルの記述をラベル名指定に変更します。**

修正前

```
/dev/sdb1 /this_year ext3 defaults 1 2
```

修正後

```
LABEL=DATA_2010 /this_year ext3 defaults 1 2
```

● swap 領域の場合

以下の手順では、＜ label ＞を「SWAP-sdc1」、＜ device ＞を「/dev/sdc1」にした場合を例に説明します。

1 **以下のコマンドを実行し、ラベル名を設定します。**

```
# /sbin/mkswap -L SWAP-sdc1 /dev/sdc1
```

2 **以下のコマンドを実行しラベル名を確認します。**

```
# /sbin/blkid -c /dev/null /dev/sdc1
```

手順 1 で設定したラベル名「/dev/sdc1: TYPE="swap" LABEL="SWAP-sdc1"」が、表示されます。

3 /etc/fstab **ファイルの記述をラベル名指定に変更します。**

修正前

```
/dev/sdc1 swap swap defaults 0 0
```

修正後

```
LABEL=SWAP-sdc1 swap swap defaults 0 0
```

■ ラベル名を変更または解除する場合の注意点

すでに設定されているラベル名を変更または解除すると、システムが起動できなくなる場合があります。
基本的には、現在設定されているラベル名で不都合がない限り、ファイルシステムのラベル名の変更や解除は行わないでください。
やむを得ずラベル名の変更、または解除が必要な場合の手順を以下に示します。

注 意

- **ファイルシステムのラベル名を変更または解除する場合は、ファイルシステムがアンマウントされた状態で行ってください。**

● ラベル名を変更する場合

1 **「■ ラベル名を設定する」(→ P.64) を参照し、新しいラベル名を設定します。**

2 **変更前のラベル名を使用している箇所を、変更後のラベル名ですべて置き換えます。**

システムのインストール時に、標準でラベル名を使用しているファイルは以下のとおりです。

```
/etc/fstab
/boot/grub/grub.conf
```

参 考

ラベル名が設定されているかどうか確認する方法

- **ファイルシステム、swap 領域に応じて、それぞれ以下の手順 2 を実行してください。**
  - 「● ファイルシステムが ext2、ext3 の場合」(→ P.65)
  - 「● swap 領域の場合」(→ P.65)

  コマンド実行後に、何も表示されない(空行のみ表示される)場合は、ラベル名が設定されていません。

● ラベル名を解除する場合

1 **以下のコマンドを実行します。**

ファイルシステム、swap 領域によって設定コマンドが異なります。

ファイルシステムが ext2, ext3 の場合

```
# /sbin/e2label <device> ""
```

swap 領域の場合

```
# /sbin/mkswap -L "" <device>
```

2　**解除前のラベル名を使用している箇所を修正します。**

システムのインストール時に、標準でラベル名を使用しているファイルは以下のとおりです。

```
/etc/fstab
/boot/grub/grub.conf
```

■ ラベル指定で対処できないシステム

ファイルシステムを使用しない以下のようなシステム設定（/etc/fstab ファイルに記述されないもの）には、ファイルシステムラベルは利用できません。

- raw デバイスを使用するシステム
- デバイス名を明に指定するミドルウェアやアプリケーションを使用するシステム

これらについては、「4.2.3 ディスク系デバイスの対処方法（udev 機能）」（→ P.67）などの他の対処法を検討してください。

4.2.3　ディスク系デバイスの対処方法（udev 機能）

Linux では /dev/sda などの従来のデバイス名とは別に /dev/disk から始まるデバイス名を割り当てることができます（udev 機能）。
ラベル指定で対処できないシステムについては、udev 機能を利用することで、デバイス名ずれを予防できます。udev 機能は以下のようなデバイス名を割り当てます。

表 4 - 3：udev 機能によるデバイス名の割り当て

デバイス名	本書での表記	説明
/dev/disk/by-path/xxxxxxxx	by-path 名 （→ P.68）	/dev/disk/by-path から始まるデバイス名 xxxxxxxx はディスクの位置情報から生成される情報
/dev/disk/by-id/yyyyyyyy	by-id 名 （→ P.68）	/dev/disk/by-id から始まるデバイス名 yyyyyyyy はディスクに設定されている識別情報から生成される情報
/dev/disk/by-uuid/zzzzzzzz	by-uuid 名	/dev/disk/by-uuid から始まるデバイス名 zzzzzzzz はディスク上のファイルシステムに設定されている識別情報から生成される情報 この機能の使用は推奨しません。
/dev/disk/by-label/wwwwwwww	by-label 名	/dev/disk/by-label から始まるデバイス名 wwwwwwww はディスクに設定されているディスクラベルから生成される情報 この機能の使用は推奨しません。
/dev/sda など	互換デバイス名	従来のデバイス名

注意

- **udev 機能で割り当てられるデバイス名は、udev 機能が開始する以前には使用することができません。詳細は、「■ udev 機能による by-id 名、by-path 名を使用できない箇所」（→ P.70）を参照してください。**

■ by-path 名

by-path 名はデバイスが接続されている位置情報から生成されるデバイス名です。デバイスの認識順のずれが発生した場合、互換デバイス名は変更されますが、by-path 名は変更されません。そのため常に同じデバイス名でデバイスを利用できます。
ただし、接続する場所を変更した場合、by-path 名は変更されるので注意が必要です。
また、ディスク関連デバイスの交換・増減設時に by-path 名が変更される場合があります。
詳細については「4.3.1 ディスク関連デバイスの交換・増減設時の注意」（→ P.72）を参照してください。

■ by-id 名

by-id 名はハードディスクに設定されているユニークな識別情報（シリアル番号など）から生成されるデバイス名です。デバイスの認識順序のずれが発生した場合、互換デバイス名は変更されますが、by-id 名は変更されません。そのため常に同じデバイス名でデバイスを利用できます。
ただし、ディスクを交換した場合は by-id 名が変更されるので注意が必要です。デバイスが by-id 名に対応していなかった場合、システム起動時にメッセージが出力されます。このメッセージはシステムに影響しませんが、メッセージを抑止したい場合は、以下の「scsi_id コマンドのメッセージを抑止する方法」（→ P.69）を参照してください。

by-id 名と比較して、「■ by-path 名」（→ P.68）のほうが故障箇所の特定などディスクの保守性に優れているため、by-path 名の指定を推奨します。

注意

ETERNUS のディスクについては、by-id 名を指定してください。

- **ファイバーチャネルスイッチに接続されたデバイスは** by-path **名が変更されてしまうことがあります。** by-id **名は不変のため、**ETERNUS **のディスクについては、**by-id **名を指定します。**

参 考

scsi_id コマンドのメッセージを抑止する方法

- **ディスクが by-id 名に対応していない場合、by-id 作成時に以下のようなメッセージが出力されます。**

```
scsi_id[4081]: 0:0:1:0: sg_io failed status 0x8 0x0 0x0 0x2
```

メッセージの内容をもとに、メッセージに対応するデバイスの vendor 名と model 名を調べてください。
以下に、/proc/scsi/scsi ファイルを参照して vendor 名と model 名を調べる例を示します。
メッセージ中の "0:0:1:0" がそれぞれ順に、"scsi:Channel:id:Lun" と対応しています。

```
# cat /proc/scsi/scsi
Host: scsi0 Channel: 00 Id: 01 Lun: 00
  Vendor: MegaRAID Model: LD0 RAID0 69400R Rev: 1FJ0
  Type: Direct-Access          ANSI SCSI revision: 02
・・・
```

/etc/scsi_id.config ファイルの末尾に以下の行を追加してください。

```
Vendor=<vendor名>, model=<model名>, options=-b
```

vendor 名が "MegaRAID"、model 名が "LD0 RAID0 69400R" の場合の例を以下に示します。

```
Vendor="MegaRAID", model="LD0 RAID0 69400R", options=-b
```

■ by-path 名、by-id 名の使用方法

デバイス名を明確に指定するコマンド、ファイル、ミドルウェアやアプリケーションに、互換デバイス名の代わりに by-path 名、by-id 名を指定できます。
デバイス名は長くなりますが、by-path 名、by-id 名の使い方は互換デバイス名と変わりはありません。

例：by-path 名の使用

```
# mount -t ext3 /dev/disk/by-path/<ディスクの位置情報から生成される情報>
/this_year
```

● 互換デバイス名から by-path 名、by-id 名を求める方法

udevinfo コマンドで互換デバイス名に対する by-path 名を表示できます。
以下に /dev/sda の by-path 名、by-id 名を表示する例を示します。

```
# udevinfo -q symlink -n /dev/sda
disk/by-path/<ディスクの位置情報から生成される情報>
disk/by-id/scsi-<ディスクの固有情報>
```

● by-path 名から互換デバイス名を求める方法

by-path 名は対応する互換デバイス名へのシンボリックリンクであり、ls -l コマンドで確認できます。以下に表示例を示します。

```
# ls -l /dev/disk/by-path/<ディスクの位置情報から生成される情報>
lrwxrwxrwx 1 root root 9 Mar 11 2005
/dev/disk/by-path/<ディスクの位置情報から生成される情報> -> ../../sda
```

● by-id 名から互換デバイス名を求める方法

by-id 名は対応する互換デバイス名へのシンボリックリンクであり、ls -l コマンドで確認できます。以下に表示例を示します。

```
# ls -l /dev/disk/by-id/scsi-<ディスクの固有情報>
lrwxrwxrwx 1 root root 9 Mar 11 2005
/dev/disk/by-id/scsi-<ディスクの固有情報> -> ../../sda
```

■ udev 機能による by-id 名、by-path 名を使用できない箇所

udev 機能は、システム起動時の早い段階で実行されデバイス名の割り当てが行われるので、通常は問題なく by-id 名や by-path 名が利用できます。しかし、これは udev 機能が実行される前には、by-id 名や by-path 名が使用できないことを意味します。

表 4 - 4 に、ディスクデバイスの指定において、by-id 名、by-path 名が使用できない箇所とその対処方法を示します。その箇所で by-id 名または by-path 名を使用した場合、システムが起動できなくなるので注意してください。

表 4 - 4：by-id 名、by-path 名が使用できない箇所と対処方法

by-id 名、by-path 名が使用できない箇所	設定ファイル	対処方法
ブートローダに対するルートファイルシステム（/）の指定	/boot/grub/grub.conf	ラベル名を使用してデバイスを指定します。
システム起動時にマウントするルートファイルシステム（/）および swap デバイスの指定 [注 1]	/etc/fstab [注 1]	ラベル名を使用してデバイスを指定します。

[注 1]：カーネルパッケージがアップデートされた場合、/etc/fstab に記述されたルートファイルシステムおよび swap デバイスのデバイス名が、初期 RAM ディスクに設定されます。このため、システム起動失敗の間接的原因になります。

4.2.4 ディスク系デバイスの対処方法（ミドルウェア機能）

ディスク系デバイスの故障によるデバイス名ずれの発生の影響をあらかじめ想定して、ミドルウェア製品を導入することも、デバイス名ずれ対策の選択肢の1つとなります。
富士通では、以下のミドルウェア製品を提供しています。機能の詳細は、各製品のマニュアルを参照してください。

■ PRIMECLUSTER GDS

システムやデータが格納されているディスク装置をミラーリングし、ディスク故障などの不測の事態からお客様の資産を守るボリューム管理ソフトウェアです。
システム構築時と異なるデバイス名でハードディスクがマウントされることを防止（ディスク切離し）する機能を有し、データ破壊を防止します。

4.2.5 ネットワーク系デバイスの対処方法

ネットワーク系デバイスは、デバイスごとに割り当てられている一意の情報（ハードウェアアドレス）を利用することで対処することが可能です。
なお、システムインストール時に接続されているすべてのネットワーク系デバイスについては、インストール処理内で自動的に対処されるため、ユーザーによる対処は不要です。

4.3 運用・保守時の考慮

運用時、保守時に考慮すべき点を説明します。それぞれ該当する箇所を参照してください。

- 「4.3.1 ディスク関連デバイスの交換・増減設時の注意」（→ P.72）
- 「4.3.2 ネットワーク系デバイスの交換・増設時の注意」（→ P.78）
- 「4.3.3 レスキューモードで起動した場合の注意」（→ P.87）

4.3.1 ディスク関連デバイスの交換・増減設時の注意

ディスク関連デバイスとは、ハードディスク装置の他、ハードディスクが接続されるカード類（SAS カード、FC カードなど）、それらのカードを収納する装置（PRIMEQUEST 1000 シリーズの SASU など）の総称です。
ディスク関連デバイスの交換・増減設時に、ディスクの by-path 名が変更されることがあります。オペレーションごとの対処方法について表 4 - 5 に示します。

表 4 - 5：by-path 名が変更されるデバイスとオペレーション、および対処方法

対象デバイス [注 1]	オペレーション	対処方法
・内蔵 SAS ディスク ・SSD ディスク ・SAN ディスク	ディスクの活性交換 [注 2]	「■ ディスクの活性交換時の by-path 名再設定手順」（→ P.73）
	ディスクの非活性交換	「■ ディスクの非活性交換時の by-path 名再設定手順」（→ P.74）
・FC カード ・SAS アレイコントローラカード ・GSPB ・SASU	左記のディスク関連デバイスの交換	「■ ディスク関連デバイス（ディスク以外）交換時の by-path 名再設定手順」（→ P.76）
・SB ・IOB ・GSPB ・SASU ・PCI Box	左記のディスク関連デバイスの増減設	「■ ディスク関連デバイス（ディスク以外）増減設時の by-path 名再設定手順」（→ P.77）

［注 1］：機種によっては、サポートされていないハードウェアがあります。
［注 2］：機種によっては、活性交換ができないものもあります。

対処方法では、/mnt にマウントするデバイスの by-path 名が /etc/fstab ファイルに記述されている場合を例に、それぞれの交換手順を示します。
コマンドの引数、システムの設定ファイル、またはミドルウェアの設定ファイルなどに by-path 名を使用している場合、ディスク関連デバイスの交換により by-path 名が変更されると、ディスクを正しく参照できなくなります。そのため、システムが正しく動作しない場合があります。
ディスク関連デバイスの交換後、ディスクの by-path 名を確認し、変更前の by-path 名を使用している箇所を変更後の by-path 名ですべて置き換えてください。

■ ディスクの活性交換時の by-path 名再設定手順

システムの電源を停止せずにディスクの交換を行う場合（活性交換）の、by-path 名の再設定手順を示します。

参 考

- **以下の手順におけるコマンド出力例は、SAS ディスクを例としています。SAS ディスク以外の活性交換可能なディスクをサポートする機種においては、出力例を読み替えてください。**

1 ディスクを活性交換する前の by-path 名を記録します。

手順 4 を実施する際に、ディスクを活性交換する前の by-path 名が必要です。/dev/disk/by-path ディレクトリ配下に対して、ls コマンドで参照した結果を記録します。

例：ディスク交換前の by-path 名を by-path-before.log に記録する

```
# ls /dev/disk/by-path/ | sort > /tmp/by-path-before.log
```

2 交換するディスクを利用しているサービスの停止や、ディスクのアンマウントなどを行います。

例：ディスクのアンマウント

```
# umount /mnt
```

3 ディスクを活性交換します。

ディスクの活性交換手順の詳細は、ご利用の機種のハードウェアマニュアルを参照してください。

4 ディスク交換前とディスク交換後の by-path 名を確認します。

/dev/disk/by-path ディレクトリ配下を ls コマンドで参照した結果を、手順 1 で記録したディスク交換前の by-path 名と比較し、ディスク交換前後の by-path 名を確認します。

例：ディスク交換前の by-path 名を確認

```
# cat /tmp/by-path-before.log
＜ディスクの位置情報から生成される情報①＞
＜ディスクの位置情報から生成される情報①＞-part1
＜ディスクの位置情報から生成される情報①＞-part2
…
```

この例では、「/dev/disk/by-path/ ＜ディスクの位置情報から生成される情報①＞」がディスク交換前の SAS ディスクの by-path 名に該当します。

例：ディスク交換後の by-path 名を確認

```
# ls /dev/disk/by-path/
＜ディスクの位置情報から生成される情報②＞
＜ディスクの位置情報から生成される情報②＞-part1
＜ディスクの位置情報から生成される情報②＞-part2
…
```

この例では、「/dev/disk/by-path/ ＜ディスクの位置情報から生成される情報②＞」がディスク交換後の SAS ディスクの by-path 名に該当します。

5　**必要に応じてパーティションの再作成などを行います。**

6　**手順 4 で確認したディスク交換前の by-path 名を使用している箇所を、ディスク交換後の by-path 名ですべて置き換えます。**

ディスクのパーティションを使用している場合は、「-part*n*」（*n* はパーティション番号）を追加してください。

例：/etc/fstab の設定を編集する

編集前

```
# cat /etc/fstab
/dev/disk/by-path/<ディスクの位置情報から生成される情報①>-part1 /mnt ext2
defaults 0 0
…
```

編集後

```
# cat /etc/fstab
/dev/disk/by-path/<ディスクの位置情報から生成される情報②>-part1 /mnt ext2
defaults 0 0
…
```

7　**手順 2 で停止したサービスの再開や、ディスクのマウントなどを行います。**

例：ディスクのマウント

```
# mount /mnt
```

■ ディスクの非活性交換時の by-path 名再設定手順

システムの電源を停止してディスクの交換を行う場合（非活性交換）の、by-path 名の再設定手順を示します。

参 考

- **以下の手順におけるコマンド出力例は、SAS ディスクを例としています。SAS ディスク以外の活性交換可能なディスクをサポートする機種においては、出力例を読み替えてください。**

1　**ディスク交換前の by-path 名を記録します。**

手順 3 を実施する際に、ディスク交換前の by-path 名が必要です。/dev/disk/by-path ディレクトリ配下に対して、ls コマンドで参照した結果を記録します。

例：ディスク交換前の by-path 名を by-path-before.log に記録する

```
# ls /dev/disk/by-path/ | sort > /tmp/by-path-before.log
```

2　**ディスクの非活性交換後、ランレベル 1（シングルユーザーモード）でシステムを起動します。**

3　ディスク交換前とディスク交換後の by-path 名を確認します。

/dev/disk/by-path ディレクトリ配下を ls コマンドで参照した結果を、手順 1 で記録したディスク交換前の by-path 名と比較し、ディスク交換前後の by-path 名を確認します。

例：ディスク交換前の by-path 名を確認

```
# cat /tmp/by-path-before.log
<ディスクの位置情報から生成される情報①>
<ディスクの位置情報から生成される情報①>-part1
<ディスクの位置情報から生成される情報①>-part2
…
```

この例では、「/dev/disk/by-path/ <ディスクの位置情報から生成される情報①>」がディスク交換前のディスクの by-path 名に該当します。

例：ディスク交換後の by-path 名を確認

```
# ls /dev/disk/by-path/
<ディスクの位置情報から生成される情報②>
<ディスクの位置情報から生成される情報②>-part1
<ディスクの位置情報から生成される情報②>-part2
…
```

この例では、「/dev/disk/by-path/ <ディスクの位置情報から生成される情報②>」がディスク交換後のディスクの by-path 名に該当します。

4　必要に応じてパーティションの再作成などを行います。

5　手順 3 で確認したディスク交換前の by-path 名を使用している箇所を、ディスク交換後の by-path 名ですべて置き換えます。

ディスクのパーティションを使用している場合は、「-part*n*」（*n* はパーティション番号）を追加してください。

例：/etc/fstab の設定を編集する

編集前

```
# cat /etc/fstab
/dev/disk/by-path/<ディスクの位置情報から生成される情報①>-part1 /mnt ext2
defaults 0 0
…
```

編集後

```
# cat /etc/fstab
/dev/disk/by-path/<ディスクの位置情報から生成される情報②>-part1 /mnt ext2
defaults 0 0
…
```

6　システムを再起動します。

```
# shutdown -r now
```

■ ディスク関連デバイス（ディスク以外）交換時の by-path 名再設定手順

システムの電源を停止してディスク以外のディスク関連デバイス（SAS カード、SATA カードなど）の交換を行う場合の、by-path 名の再設定手順を示します。

参 考

- **以下の手順におけるコマンド出力例は、SAS カードを例としています。SAS カード以外のディスク関連デバイスをサポートする機種においては、出力例を読み替えてください。**

1. **ディスク関連デバイス（ディスク以外）の交換後、ランレベル 1（シングルユーザーモード）でシステムを起動します。**

2. **by-path 名を使用している箇所を見て、変更前の by-path 名を確認します。**

   /dev/disk/by-path/＜ディスクの位置情報から生成される情報＞は、以下の形式で出力されます。
   形式：/dev/disk/by-path/pci-xxxx:xx:xx.x-sas-w:w:w-w:w

   ```
   # cat /etc/fstab
   …
   /dev/disk/by-path/pci-xxxx:xx:xx.x-sas-w:w:w-w:w-part1 /mnt ext2
   defaults 0 0
   …
   ```

   この例では、「/dev/disk/by-path/pci-xxxx:xx:xx.x-sas-w:w:w-w:w-part1」が変更前の by-path 名に該当します。

3. **/dev/disk/by-path ディレクトリ配下を ls コマンドなどで参照した結果を、手順 2 で確認した変更前の by-path 名と比較し、変更後の by-path 名を確認します。**

   交換後の by-path 名は、交換前の by-path 名から以下に示した下線部分だけが変更されます。
   /dev/disk/by-path/pci-xxxx:xx:xx.x-sas-z:z:z-w:w-part1

   ```
   # ls /dev/disk/by-path/
   pci-xxxx:xx:xx.x-sas-z:z:z-w:w
   pci-xxxx:xx:xx.x-sas-z:z:z-w:w-part1
   pci-xxxx:xx:xx.x-sas-z:z:z-w:w-part2
   pci-xxxx:xx:xx.x-sas-z:z:z-y:y
   pci-xxxx:xx:xx.x-sas-z:z:z-y:y-part1
   …
   ```

   この例では、「/dev/disk/by-path/pci-xxxx:xx:xx.x-sas-z:z:z-w:w-part1」が変更後の by-path 名に該当します。

4. **変更前の by-path 名を使用している箇所を、手順 3 で確認した変更後の by-path 名ですべて置き換えます。**

5　**システムを再起動します。**

```
# shutdown -r now
```

注 意

- **PCI スロットからカードを取り外した場合は、デバイス名ずれが発生するため、一度搭載したカードは原則削除しないでください。**

■ ディスク関連デバイス（ディスク以外）増減設時の by-path 名再設定手順

システムの電源を停止してディスク関連デバイス（ディスク以外）の増設、取り外しを行う場合の、by-path 名の再設定手順を示します。

参 考

- **以下の手順におけるコマンド出力例は、SAS ディスクを例としています。SAS ディスク以外のディスク関連デバイスをサポートする機種においては、出力例を読み替えてください。**

1　**ディスク関連デバイス（ディスク以外）の増設または取り外し後、ランレベル 1（シングルユーザーモード）でシステムを起動します。**

2　**by-path 名を使用している箇所を見て、変更前の by-path 名を確認します。**

/dev/disk/by-path/ <ディスクの位置情報から生成される情報>は、以下の形式で出力されます。
形式：/dev/disk/by-path/pci-xxxx:xx:xx.x-sas-w:w:w-w:w

```
# cat /etc/fstab
…
/dev/disk/by-path/pci-xxxx:xx:xx.x-sas-w:w:w-w:w-part1 /mnt ext2
defaults 0 0
…
```

この例では、「/dev/disk/by-path/pci-xxxx:xx:xx.x-sas-w:w:w-w:w-part1」が変更前の by-path 名に該当します。

3 **/dev/disk/by-path ディレクトリ配下を ls コマンドなどで参照した結果を、手順 2 で確認した変更前の by-path 名と比較し、変更後の by-path 名を確認します。**

交換後の by-path 名は、交換前の by-path 名から以下に示した下線部分だけが変更されます。

/dev/disk/by-path/pci-xxxx:xx:xx.x-sas-w:w:w-w:w-part1

```
# ls /dev/disk/by-path/
pci-zzzz:zz:zz.z-sas-w:w:w-w:w
pci-zzzz:zz:zz.z-sas-w:w:w-w:w-part1
pci-zzzz:zz:zz.z-sas-w:w:w-w:w-part2
pci-zzzz:zz:zz.z-sas-y:y:y-y:y
pci-zzzz:zz:zz.z-sas-y:y:y-y:y-part1
…
```

この例では、「/dev/disk/by-path/pci-zzzz:zz:zz.z-sas-w:w:w-w:w-part1」が変更後の by-path 名に該当します。

4 **変更前の by-path 名を使用している箇所を、手順 3 で確認した変更後の by-path 名ですべて置き換えます。**

5 **システムを再起動します。**

```
# shutdown -r now
```

4.3.2 ネットワーク系デバイスの交換・増設時の注意

ネットワーク系デバイスは、/etc/sysconfig/network-scripts ディレクトリ配下の ifcfg-ethX ファイルで管理されます。「ifcfg-ethX」の "ethX" がネットワークデバイス名を表します。"X" は整数で、ifcfg-eth0、ifcfg-eth1 のようなファイル名になります。
このファイルにネットワーク系デバイスのハードウェアアドレス（MAC アドレス）を定義すると、ネットワークデバイス名を固定にすることができます。
具体的には、ifcfg-ethX ファイルに以下の行を追加します。

```
HWADDR=<MAC-address>
```

＜*MAC-address* ＞は、AA:BB:CC:DD:EE:FF 形式でイーサネットデバイスのハードウェアアドレスを指定します。
イーサネットデバイスのハードウェアアドレスは、/etc/sysconfig/hwconf ファイルを参照したり、ifconfig コマンドを実行したりすることで確認できます。
以降の記述は、bonding を構成する slave デバイスにも適用可能です。

■ ネットワークデバイス名の固定手順

1　ハードウェアアドレス（MAC アドレス）を確認します。

例：/etc/sysconfig/hwconf ファイルを参照した場合

```
# grep -A 5 eth0 /etc/sysconfig/hwconf
device: eth0
driver: e1000
desc: "Intel Corp.|82540EM Gigabit Ethernet Controller"
network.hwaddr: 00:0A:E4:06:39:A1
vendorId: 8086
```

例：ifconfig コマンドを実行した場合

```
# ifconfig
eth0 Link encap:Ethernet HWaddr 00:0A:E4:06:39:A1
```

ハードウェアアドレス「00:0A:E4:06:39:A1」が確認できます。

2　/etc/sysconfig/network-scripts ディレクトリ配下の ifcfg-eth*X* ファイルに、手順 1 で確認したハードウェアアドレス（MAC アドレス）を定義します。

例 : /etc/sysconfig/network-scripts/ifcfg-eth0 ファイル

```
DEVICE=eth0
BOOTPROTO=static
BROADCAST=192.168.1.255
HWADDR=00:0A:E4:06:39:A1
IPADDR=192.168.1.110
NETMASK=255.255.255.0
NETWORK=192.168.1.0
ONBOOT=yes
TYPE=Ethernet
```

■ NIC の増設、取り外し、交換の手順

以下の作業では、途中でシステムを再起動します。このとき、通常モード、シングルユーザーモードのいずれを用いても構いませんが、モードによって、手順が異なる場合があります。

● NIC の増設

増設にあたり、サーバの電源断前後でそれぞれ作業があります。

1　存在するネットワークデバイス名を確認します。

/etc/sysconfig/network-scripts ディレクトリ配下に存在する ifcfg-ethX ファイル名を参照し、現在システムに存在するネットワークデバイス名（ethX）を確認します。

2　NIC を増設します。

サーバの電源断後、ご利用の機種のハードウェアマニュアルに記載されている、オプションカードの追加／取り外し手順を参照し、NIC を増設してください。NIC の増設後、サーバの電源を入れてください（システムを再起動します）。

3 modprobe.conf ファイルを編集します。

/etc/modprobe.conf ファイルに追加した NIC のエントリーがない場合には追加してください。

```
alias  ethX e1000      //e1000の場合
```

4 追加されたネットワークデバイス名を確認します。

通常モードでシステムを起動した場合は、増設したネットワークデバイス用の設定ファイルが /etc/sysconfig/network-scripts ディレクトリ配下に作成されます。このディレクトリ配下に存在する ifcfg-ethX ファイル名を参照し、手順 1 の時点で存在していなかったネットワークデバイス名（ethX）を確認します。
シングルユーザーモードでシステムを起動した場合は、上記ファイルが作成されないので、以下のコマンドを実行してファイルを作成します。

```
# kudzu -q
```

ここで確認したネットワークデバイス名が、増設された NIC に搭載されているネットワークデバイスに該当します。

5 ifcfg-ethX ファイルを編集します。

必要に応じ、追加したネットワークデバイスの設定ファイル（ifcfg-ethX ファイル）を編集してください。bonding デバイスの場合、MASTER オペランド、SLAVE オペランドなどを追加します。なお、ネットワークデバイス名を固定するためのハードウェアアドレスは自動的に設定されるので、ユーザーによる対処は不要です。
以下に、自動的に作成される設定ファイルの例を示します。

```
# Intel Corporation 82545GM Gigabit Ethernet Controller
DEVICE=eth3
BOOTPROTO=dhcp
ONBOOT=yes
HWADDR=00:04:23:b2:12:b7
TYPE=Ethernet
```

6 ネットワークデバイスを有効にします。

手順 5 で設定ファイルを編集した場合

編集内容を反映するために、以下のコマンドを実行します。

```
# /sbin/ifdown ethX
# /sbin/ifup ethX
```

手順 5 で設定ファイルを編集しなかった場合

すでにネットワークデバイスは有効になっているため、操作は不要です。

シングルユーザーモードでシステムを起動した場合は、設定ファイルの編集の有無にかかわらず、通常運用に復帰するため、再起動が必要です（ifdown/ifup は不要です）。

注 意

- **bonding デバイス自体を追加する場合は、いったん slave デバイスを通常のデバイスとして追加し、改めて設定してください。**

● NIC の取り外し

カードを取り外した場合は、増設と逆の現象となりデバイス名が変更されるため、一度搭載したカードは原則取り外さないでください。
やむを得ず取り外す場合の手順を以下に示します。

1 ifcfg-ethX ファイルを削除します。

/etc/sysconfig/network-scripts ディレクトリ配下に存在する、これから取り外す NIC に搭載されているネットワークデバイスに対応する設定ファイルを削除します。

2 NIC を取り外します。

サーバの電源断後、ご利用の機種のハードウェアマニュアルに記載されている、オプションカードの追加／取り外し手順を参照し、NIC を取り外してください。

● NIC の交換

複数のカードを一度に交換するとデバイス名が変更される場合があるため、カードの交換は 1 枚ずつ実施してください。

1 ネットワークデバイス名とバスアドレスの対応一覧を作成します。

NIC の交換の場合、交換前／交換後の NIC がどのスロットに装着されているかを意識する必要があります。このため、「● ネットワークデバイス名と LAN ポートのバスアドレス・MAC アドレスの対応一覧」（→ P.84）を参照し、ネットワークデバイス名とバスアドレスの対応を表す一覧表を作成してください。

2 交換対象となる NIC のネットワークデバイスを確認します。

NIC に対応するネットワークデバイス名の確認は、NIC の故障の場合、それにより通信が不可能となるネットワークデバイス名を確認することで行われます。

注 意

- **NIC が複数 LAN ポートを有している場合、故障していないネットワークデバイスも、NIC の交換により影響を受けます。あらかじめ交換前に、影響を受けるネットワークデバイス名を特定してください。特定の仕方については、「● 同一カード上のネットワークデバイスの特定方法」（→ P.86）を参照してください。**

3 交換対象となるネットワークデバイスの ifcfg-ethX ファイルを退避します。

```
# cd /etc/sysconfig/network-scripts
# mkdir temp
# mv ifcfg-ethX  temp/ifcfg-ethX.bak
```

注 意

- **NIC が複数 LAN ポートを有している場合､影響を受けるネットワークデバイスすべてに対して､上記操作を実施してください。**
- **ファイルの退避は、必ずサブディレクトリを作成して行ってください。システムの起動時に /etc/sysconfig/network-scripts 配下のファイルがスキャンされるため、ファイル名変更だけでは退避したことにならず、システムに影響する場合があります。**

4 modprobe.conf ファイルを退避します。

通常、このファイルが破壊される危険性はありませんが、念のため退避しておきます。

```
# mkdir /etc/bak
# cp /etc/modprobe.conf  /etc/bak/modprobe.conf
```

5 システムの電源を切断し、NIC を交換します。

サーバの電源断後、ご利用の機種のハードウェアマニュアルに記載されている、オプションカードの追加／取り外し手順を参照し、NIC を交換してください。NIC の交換後、サーバの電源を入れてください（システムを再起動します）。

6 ifcfg-ethX.bak ファイルを編集します。

手順 3 で /etc/sysconfig/network-scripts/temp 配下に退避した設定ファイル（ifcfg-ethX.bak）に、交換後の NIC の MAC アドレスを設定します。NIC の種類により以下の対応を行います。

NIC の LAN ポートが 1 つの場合

通常モードでシステムを起動した場合は、システム起動時に、交換した NIC に対するネットワークデバイス用の設定ファイルが、新たに /etc/sysconfig/network-scripts ディレクトリ配下に作成されます。新たに作成された設定ファイル（ifcfg-ethX）には、交換後の NIC の MAC アドレスが設定されています。/etc/sysconfig/network-scripts/temp 配下に退避した設定ファイル（ifcfg-ethX.bak）内に書かれている HWADDR の値を、新たに作成された設定ファイル（ifcfg-ethX）に書かれている HWADDR の値に置き換えてください。

シングルユーザーモードでシステムを起動した場合は、上記ファイルが作成されないので、以下のコマンドを実行してファイルを作成します。

```
# kudzu -q
```

その後、同様に、/etc/sysconfig/network-scripts/temp 配下に退避した設定ファイル（ifcfg-ethX.bak）を編集して MAC アドレスを反映します。

NIC の LAN ポートが複数の場合

通常モードでは、システム起動時に、交換した NIC に対する複数のネットワークデバイス用の設定ファイルが、新たに /etc/sysconfig/network-scripts ディレクトリ配下に作成されます。シングルユーザーモードでシステムを起動した場合は、上記ファイルが作成されないので、kudzu を実行してファイルを作成します。

このとき、「4.1.4 デバイス名ずれの要因：ハードウェア構成変更」（→ P.51）で説明のように、ネットワークデバイスにデバイス名ずれが発生している可能性があります。この場合、「■ 複数 LAN ポートを有する NIC を交換する場合のネットワークデバイス名の固定手順」（→ P.83）に示す方法で、各ネットワークデバイス名に対する正しい MAC アドレスを確認し、/etc/sysconfig/network-scripts/temp 配下に退避した各設定ファイル（ifcfg-ethX.bak）内の HWADDR に、MAC アドレスを設定してください。

7　**以下のコマンドを実行し、ネットワークデバイス用の設定ファイルを復元します。**

```
# cd /etc/sysconfig/network-scripts
# mv temp/ifcfg-ethX.bak  ifcfg-ethX
```

注 意

- **NIC が複数 LAN ポートを有している場合、影響を受けるネットワークデバイスすべてに対して、上記操作を実施してください。**

8　**modprobe.conf ファイルを復元します。**

手順 4 で退避した /etc/modprobe.conf ファイルを復元します。

```
# mv /etc/bak/modprobe.conf /etc/modprobe.conf
```

9　**以下のコマンドを実行し、ネットワークデバイスを有効にします。**

```
# /sbin/ifdown ethX
# /sbin/ifup ethX
```

シングルユーザーモードでシステムを起動した場合は、通常運用に復帰するため、再起動が必要です（ifdown/ifup は不要です）。

注 意

- **NIC が複数 LAN ポートを有している場合、影響を受けるネットワークデバイスすべてに対して、上記操作を実施してください。**

10　**作成した退避ディレクトリを削除します。**

```
# rm -r /etc/sysconfig/network-scripts/temp
# rm -r /etc/bak
```

■ 複数 LAN ポートを有する NIC を交換する場合のネットワークデバイス名の固定手順

複数の LAN ポートを持つ NIC やシステムボードを交換する場合は、「4.1.4 デバイス名ずれの要因：ハードウェア構成変更」（→ P.51）で説明のように、交換後に起動したシステムにおいて、ネットワークデバイス名にデバイス名ずれが発生している可能性があります。
デバイス名ずれの発生の有無は、以下の手順で確認します。

1　**交換前と交換後の LAN ポートについて、バスアドレス（物理位置に相当する情報）とネットワークデバイス名の対応を確認します。**

交換前と交換後で、バスアドレスとネットワークデバイス名の対応が一致しない場合、デバイス名ずれが発生していることを意味します。

2　デバイス名ずれが発生していた場合は、次の対処を行います。

・交換前と同じバスアドレスの LAN ポートの MAC アドレスを確認します。
・確認した MAC アドレスを、ifcfg-ethX ファイルへ登録します。

この対応を行うには、交換前、交換後、それぞれのネットワークデバイス名と LAN ポートのバスアドレス・MAC アドレスの対応一覧が必要です。以下にネットワークデバイス名と LAN ポートのバスアドレス・MAC アドレスの対応一覧の作成手順と、デバイス名ずれ発生の有無の確認およびデバイス名固定手順について説明します。

● ネットワークデバイス名と LAN ポートのバスアドレス・MAC アドレスの対応一覧

LAN ポートのバスアドレスは、データ・バス上の以下の情報の組み合わせで示されます。

<バス番号>*: *<スロット番号>*: *<ファンクション番号>

稼働中のシステムにおいて、ネットワークデバイス名 (ethX) と対応する LAN ポートのバスアドレス・MAC アドレスの確認方法を以下に示します。この方法を使用してネットワークデバイス名とバスアドレス・MAC アドレスの対応一覧を作成してください。

1　/sys/class/net ディレクトリ配下に移動し、ネットワークデバイス名に相当するディレクトリを確認します。

```
# cd /sys/class/net
# ls
eth0  eth1  lo
```

2　ネットワークデバイス名をもとに、以下の操作を行います。

eth0 デバイスを例に示します。

1. **ネットワークデバイス名とバスアドレスの対応を確認します。**

```
# ls -l eth0/device
lrwxrwxrwx 1 root root 0 Apr  9 09:17 eth0/device
-> ../../../devices/pci0000:01/0000:01:06.0
```

上記出力結果から以下を確認し、ネットワークデバイス名とバスアドレス（*<バス番号>*: *<スロット番号>*: *<ファンクション番号>*）の対応一覧を作成してください。
・ネットワークデバイス名：/sys/class/net 配下のディレクトリ名
・バスアドレス：シンボリックリンク先ファイル名のうち一番下位層のファイル名（実際はディレクトリ名）

2. **デバイス名と MAC アドレスの対応を確認します。**

```
# cat eth0/address
00:30:05:30:ef:8d
```

出力結果は、そのデバイス名の MAC アドレスとなります。
対象となるデバイスを、bonding デバイスの slave デバイスとして組み込む場合は、bonding の設定をする前に作業を実施するか、シングルユーザーモードで起動して作業してください。

3 /sys/class/net ディレクトリ配下にあるすべての ethX ディレクトリに対して、手順 2 を実施します。

参 考

- ネットワークデバイス名とバスアドレス・MAC アドレスの対応一覧は、以下の例に示す表形式にまとめて管理することを推奨します。

表 4 - 6：ネットワークデバイス名とバスアドレス・MAC アドレス対応一覧 (例)

作成日	yyyy/mm/dd	
対象システム	＜サーバ名＞	
ネットワークデバイス名	バスアドレス（バス : スロット : ファンクション）	MAC アドレス
eth0	1:6:0	00:30:05:30:ef:8d
eth1	1:6:1	00:30:05:30:ef:8e
…	…	…

● ネットワークデバイスのデバイス名ずれ発生の確認とデバイス名固定手順

複数の LAN ポートを持つ NIC やシステムボードを交換した場合の、デバイス名ずれ発生の確認手順と、デバイス名の固定手順を説明します。

注 意

- ハードウェア故障による予期しない突発的な交換作業の発生に備えて、システム導入時には、本作業を実施することを推奨します。

1 NIC またはシステムボードの交換前に、ネットワークデバイス名とバスアドレス・MAC アドレスの対応一覧を作成します。
「● ネットワークデバイス名と LAN ポートのバスアドレス・MAC アドレスの対応一覧」（→ P.84）を参照して作成してください。

2 システムを停止し、NIC またはシステムボードを交換します。

3 システムを起動します。

4 交換後のネットワークデバイス名とバスアドレス・MAC アドレスの対応一覧を作成します。
「● ネットワークデバイス名と LAN ポートのバスアドレス・MAC アドレスの対応一覧」（→ P.84）を参照して作成してください。

5 手順 1 で作成した一覧表と、手順 4 で作成した一覧表を比較し、内容が一致しないデバイスを抽出します。

参 考

- **NIC またはシステムボードが交換されると、新しい MAC アドレスが割り振られるため、交換により影響を受けたデバイスの内容は、交換前後で必ず不一致となります。**

6 手順 5 で確認された内容が一致しないデバイスに対して、「デバイス名－バスアドレス」の対応状況を交換前後で比較します。

交換前後で「デバイス名－バスアドレス」の対応が一致している場合

デバイス名ずれは発生していません。
この場合、交換後の一覧にある MAC アドレスを、そのまま ifcfg-ethX に登録します。

交換前後で「デバイス名－バスアドレス」の対応に不一致があった場合

デバイス名ずれが発生しています。以下の対処を行ってください。

1. **対象とするネットワークデバイス名の交換前のバスアドレスを確認します。**
2. **確認したバスアドレスを交換後の一覧内から確認します。**
3. **バスアドレスに対応する MAC アドレスを取得し、ifcfg-ethX ファイルへ登録します。登録方法は「■ ネットワークデバイス名の固定手順」（→ P.79）を参照してください。**
4. **システムを再起動します。**

```
# shutdown -r now
```

再起動後、交換前と同じネットワーク構成になります。

参 考

- **ネットワークインターフェースに割り当てる名前に故意に欠番を設けた場合（eth0,eth1,eth3,eth4 など）、欠番より大きな番号のインターフェース（この例では、eth3,eth4）の交換を行うと、システムはより若い空いた番号を使うため、システムの再起動時にインターフェース名が変更される場合があります。この場合、上記で説明した方法で MAC アドレスを確認し、デバイス名の固定手順を実施してください。**
  また、交換時の名前に関する混乱を避けるために、インターフェース名に欠番が生じるような設計はしないでください。

● 同一カード上のネットワークデバイスの特定方法

複数 LAN ポートを有する NIC において、1 つの LAN ポートが故障したことにより NIC を交換する場合、故障していない他の LAN ポートに対応するネットワークデバイスも、NIC 交換の影響を受けます。
このとき、影響を受ける他のネットワークデバイスを特定する方法を以下に示します。

1 NIC またはシステムボードの交換前に、ネットワークデバイス名とバスアドレス・MAC アドレスの対応一覧を作成します。

「● ネットワークデバイス名と LAN ポートのバスアドレス・MAC アドレスの対応一覧」（→ P.84）を参照して作成してください。

2 交換対象となる NIC のネットワークデバイスを確認します。

NIC の故障により通信が不可能となったネットワークデバイス名を確認してください。

3 **交換対象となるネットワークデバイスのバスアドレスを確認します。**

手順 1 で作成した一覧をもとに、ネットワークデバイスのバスアドレスを特定します。

4 **特定したバスアドレスと、同一のバス番号・スロット番号を持つネットワークデバイスを確認します。**

LAN ポートのバスアドレスは、データ・バス上の以下情報の組み合わせで示されます。

<バス番号> : ***<スロット番号>*** : ***<ファンクション番号>***

このうち、ファンクション番号のみ異なり、バス番号、スロット番号が同一のバス番号を有するネットワークデバイスは、同一の NIC 上に存在することを意味し、NIC 交換により影響を受けるデバイスとなります。

4.3.3 レスキューモードで起動した場合の注意

インストール CD のレスキューモードを利用することにより、レスキューシステムを起動し、トラブルが発生して起動しなかったシステムを復旧できる可能性があります。
しかし、レスキューモードを利用した場合、通常運用時のデバイス名とレスキューモードで使用するデバイス名が異なる場合があります。

■ ディスク系デバイス／ネットワーク系デバイス共通の注意

ドライバの組み込み順序が変わることにより、デバイス名ずれが発生する場合があります。
通常運用時は、「4.1.1 デバイスの認識方法とデバイス名の決まり方」（→ P.47）で記述している順序でデバイスを認識します。
ただし、レスキューモードでシステムを起動する際はユーザーが設定した /etc/modprobe.conf ファイルを参照しないため、ドライバの組み込み順序が異なり通常運用時とは異なる順序でデバイスを認識する場合があります。
このため、システム起動時のオプションとして noprobe または nostorage を指定して、自動でドライバが組み込まれることを抑止し、ドライバの組み込み順序を制御してください。
ドライバの組み込み順序は通常運用時の指定順と同じになるようにしてください。

■ ディスク系デバイスの注意

● USB-CDROM 使用時の注意

インストール CD を USB-CDROM（DVD-ROM&CD-R/RW ドライブユニット）に入れて noprobe または nostorage 指定でレスキューモードを起動した場合は、USB ドライバを最初に組み込みます。そのため、USB-CDROM が /dev/sda となり、以降のデバイス名がずれます。
デバイスを操作する際は、デバイス名がずれていることに注意してください。

● udev 機能によるデバイス名ずれ防止策を実施の場合の注意

「4.2.3 ディスク系デバイスの対処方法（udev 機能）」（→ P.67）は、システムディスク以外のディスクを対象としたデバイス名ずれの防止策です。しかし、レスキューモードでは udev 機能が有効とならないため、本防止策は機能しません。

したがって、レスキューモードの起動はシステムディスクの復旧を行う際に使用するにとどめ、データ用ディスクの復旧は、システムをシングルユーザーモードで起動し、udev 機能が有効となった状態での実施を推奨します。

■ ネットワーク系デバイスの注意

レスキューモードでは、通常運用時に設定した /etc/sysconfig/network-scripts ディレクトリ配下の ifcfg-ethX ファイルは使用されません。したがって、レスキューモードの起動画面において、またはレスキューモード起動後に ifconfig コマンドを使用して、デバイス名を考慮したネットワーク設定を実施してください。

例：ifconfig コマンドによるネットワークの設定

```
# /etc/rc.d/init.d/network start
# ifconfig eth0 192.168.1.110 netmask 255.255.255.0 broadcast 192.168.1.255
```

5

第 5 章

システム設計：ログファイル運用

この章では、ログファイル運用について説明します。

5.1 ログファイル運用設計の必要性と対象ログファイル

ログファイルは、システムの保守や監視に欠かせない情報です。
ログファイルの運用を設計する必要性と、ログファイルの種類について説明します。

- 「5.1.1 ログファイル運用設計の必要性」（→ P.90）
- 「5.1.2 ログファイルの種類」（→ P.92）

5.1.1 ログファイル運用設計の必要性

ログファイル運用設計の必要性について説明します。

■ ログファイルの役割

システムやアプリケーションで発生した様々な事象は、システムやアプリケーションが作成するログファイルに記録されます。
ログファイルに記録された情報から、システムの稼働状況や、異常事象発生など、状態変化の履歴を確認できます。ログファイルは、システム保守を行う際に重要な情報となります。

■ 1 つのログファイルで運用する場合の問題点

ログファイルは通常、出力されたログを蓄積し続けます。そのままの運用を継続した場合、ログファイルはディスク領域の空き容量を消費し尽くし（ファイルシステム・フルの発生）、それ以上はログをファイルに書き込めなくなります。
これによりログのロストが発生し、異常発生時などに必要な情報を収集できなくなる可能性があります。

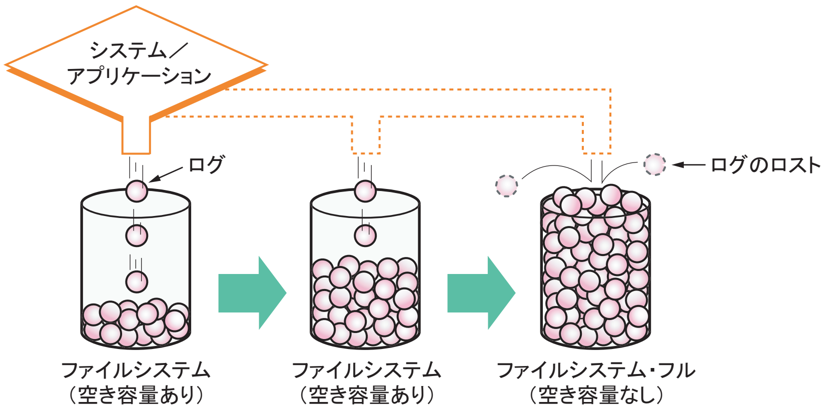


図 5 - 1：1 つのログファイルで運用する場合

ログのロストを避けるためには、ディスク領域がフルになる前に、ログファイルをバックアップ媒体に退避し、ディスク領域を解放する必要があります。

また、1 つのログファイルにログを出力し続ける運用では、ログファイルが常に使用中であることから、正確なバックアップが保障されない、ディスク領域を解放できない、という問題が発生します。

■ ログローテーション

ログをロストすることなく、ログファイルを運用する方式として、複数のログファイルをローテーションして運用する方式があります。この運用方式を「ログローテーション」といいます。
システム保守に必要なログファイルについて、ログローテーションを実施します。
ログローテーションは、以下のとおり実現されます。

1 **複数のログファイルを用意する（図 5 - 2 ファイル A、ファイル B）**

2 **1 つ目のログファイル（図 5 - 2 ファイル B）にログを蓄積する**

3 **1 つ目のログファイルが切り替え条件（ファイルサイズ、切り替え期間など）に達したら、2 つ目のログファイル（図 5 - 2 ログファイル A）に切り替える**

4 **1 つ目のログファイル（図 5 - 2 ファイル B）をバックアップしてから削除し、ディスク領域を解放する**

手順 2 ～手順 4 を繰り返すことで、ログファイルをローテーションさせます。

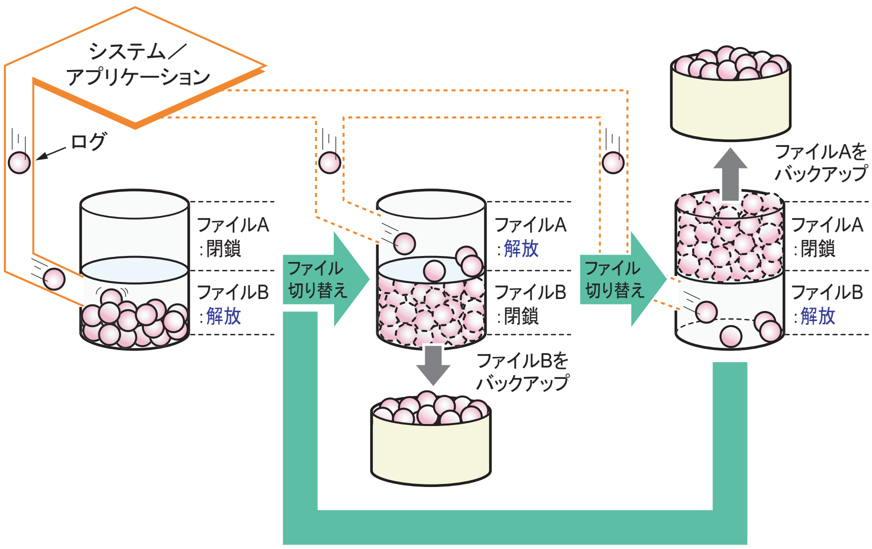


図 5 - 2：ログファイルのローテーション

5.1.2 ログファイルの種類

ログファイルには、システムログファイルと、ユーザー固有ログファイルの 2 種類があります。

■ システムログファイル

システムログファイルは表 5 - 1 に示すものです。
システムログファイルの運用設計の考え方については、「5.2 システムログファイルのローテーション運用」(→ P.93) を参照してください。

表 5 - 1：システムログファイル一覧

ログファイル名	出力プロセス	内容	切り替え契機（デフォルト）	保有世代（デフォルト）
/var/log/wtmp	各種認証プロセス（login, gdm など）	ログインの記録（ログイン種別、プロセス番号、ユーザー名、端末名、リモートホスト名、ログイン／ログアウト時刻、終了状態など）	月ごと	1
/var/log/dmesg	/etc/rc.sysinit 内の dmesg コマンド	システム起動時にコンソールに出力されるメッセージの記録（CPU 数、メモリ状況、各種ハードウェア状態など）	ローテーション運用なし boot 時に上書き	
/var/log/boot.log	syslogd	システム起動時に実行される rc スクリプトや各種サービスの記録	週ごと	4
/var/log/messages	syslogd	システム状態の記録（デバイスドライバ／カーネル／アプリケーションにより検出された、システムの状態変化やエラー事象が通知される）	週ごと	4
/var/log/cron	syslogd	クロックデーモン（cron, at）によるスケジュール実施状況の記録	週ごと	4
/var/log/secure	syslogd	認証関連の記録（ログインセッションの開始／終了や、パスワード認証の成否など）	週ごと	4

■ ユーザー固有ログファイル

ユーザー固有ログファイルは、システムログファイル以外で、ユーザー自身が必要に応じて設定しているログファイルです。
ユーザー固有ログファイルの運用設計の考え方については、「5.3 ユーザー固有ログファイルのローテーション運用」(→ P.97) を参照してください。

5.2 システムログファイルのローテーション運用

デフォルトで設定されている、システムログファイルのローテーション運用方法の見直しについて説明します。

- 「5.2.1 ローテーション運用の実施方法とスケジュール」(→ P.93)
- 「5.2.2 システムログファイルの切り替え周期と保有世代数」(→ P.94)
- 「5.2.3 システムログファイルのバックアップ」(→ P.96)

参 考

- **デフォルトの運用を変更する場合、変更方法については「5.4 システムログのローテーション運用設定」(→ P.102) を参照してください。**

5.2.1 ローテーション運用の実施方法とスケジュール

システムが出力するシステムログファイルのローテーション運用は、デフォルトでは logrotate コマンドを使用して行われます。システム設計の際、業務要件と照らし合わせて、必要に応じて見直し、変更してください。

■ 実施方法

logrotate コマンドは常駐プロセスではありません。システムログファイルのローテーション運用を行うタイミングで logrotate コマンドを起動します。
ローテーション運用を設計する際は、以下のいずれかの実施方法から検討します。

- コマンドライン入力やスクリプトから任意のタイミングで実行
- スケジュール設定による定期的な起動(cron、Systemwalker などを使用)

■ スケジュール

Linux では、cron により毎日午前 4 時 2 分に logrotate コマンドによるローテーション運用が行われるよう、デフォルトでスケジュール設定されています。
ローテーション運用は、システムログファイル出力の負荷が高い時間帯での実施は避け、業務に影響しない時間帯に実施されるよう設計を行ってください。
スケジュールの変更方法については、「5.4.1 スケジュールの変更」(→ P.102) を参照してください。

5.2.2 システムログファイルの切り替え周期と保有世代数

システムログファイルの総量がシステム運用に適切となるよう、切り替え周期および保有世代数の設計を行ってください。
切り替え周期および保有世代数の設定方法については、「5.4.2 切り替え周期などの変更」(→ P.103)を参照してください。

■ システムログファイルの切り替え周期の指定と判定条件

logrotate コマンドによるローテーション運用が実施されると、logrotate コマンドの設定ファイル（以降、設定ファイルと表記します）に記述されたシステムログファイルに対して切り替え周期に達したか否かを確認し、条件を満たしている場合にローテーション運用を実施します。
logrotate コマンドによるシステムログファイル切り替え周期の指定、および周期の指定に対する切り替え判定条件は、以下のとおりです。

表 5 - 2：logrotate コマンドのシステムログファイル切り替え周期の指定と切り替え判定条件

周期の指定	切り替え判定条件
設定ファイルにおける指定	
monthly（月ごと）	コマンド起動日が、システムログファイルの前回切り替え日と比べて、月が変わっている場合。 ・切り替わる例 　前回切り替え日が 8 月 30 日で、コマンド起動日が 9 月 1 日の場合 ・切り替わらない例 　前回切り替え日が 9 月 1 日で、コマンド起動日が 9 月 30 日の場合
weekly（週ごと）	コマンド起動日が、システムログファイルの前回切り替え日と比べて、週が変わっている場合（日曜日を週の始まりとする）。 または、前回切り替え日から 7 日以上経過している場合。 ・切り替わる例 　前回切り替え日が土曜日で、コマンド起動日が翌日の日曜日の場合 ・切り替わらない例 　前回切り替え日が日曜日で、コマンド起動日が同じ週の土曜日の場合
daily（日ごと）	コマンド起動日が、システムログファイルの前回切り替え日と異なる場合。 ・切り替わる例 　前回切り替え日が 9 月 1 日 23:59 で、コマンド起動日が 9 月 2 日 0:01 の場合 ・切り替わらない例 　前回切り替え日が 9 月 1 日 0:01 で、コマンド起動日が 9 月 1 日 23:59 の場合
size= *<ファイルサイズ>*（サイズ指定）	コマンド起動時に、システムログファイルが指定サイズを超えている場合。 ・切り替わる例 　指定サイズが 100k で、コマンド起動時にシステムログファイルのサイズが 100k 以上の場合
コマンドオプション	
強制切り替え（[-f] オプション指定）	コマンド起動時に、設定ファイルで指定のシステムログファイルすべてを無条件に切り替え。

■ システムログファイルの切り替え周期に関する注意事項

logrotate コマンドにおけるログローテーションの切り替え周期に関する注意事項は、以下のとおりです。

● システムログファイルの切り替え周期に合わせた logrotate コマンド実施タイミングの設計

設定ファイルで指定した、各システムログファイル切り替え周期の期間内で 、logrotate コマンドが実施されない場合、期待した切り替え周期でのローテーションが保証されません。
各システムログファイルの切り替え周期を考慮して、logrotate コマンド実施のタイミングの設計を行ってください。

● システムログファイルの強制切り替えを一部のシステムログファイルに対して実施する場合

logrotate コマンドに、[-f] オプションを指定して起動した場合、設定ファイルに指定されているシステムログファイルすべてに対して、強制的に切り替えが実施されます。
設定ファイル内のすべてのシステムログファイルに対し、切り替えを実施することが運用上適当ではない場合は、強制切り替えを実施するシステムログファイルだけを含む設定ファイルを個別設定ファイルとして別途用意し、通常の運用と使い分けて実施してください。

● 周期の指定期間内に 2 回以上の logrotate コマンドの実施をスケジュールする場合

周期の指定を monthly ／ weekly ／ daily に設定しているシステムログファイルは、周期の指定期間内に 1 回でも切り替わりが行われたものについては、それ以降のその期間内の切り替えは実施されませんので、注意してください。
例えば、周期の指定が daily のシステムログファイルの場合、当日に 1 回切り替わりが行われると、それ以降、その日の切り替えは実施されません。

● ファイルサイズを指定する場合

ファイルサイズの指定は、logrotate コマンドの実行時に指定サイズ以上であればローテーションが実施されることを意味しており、指定サイズ以上になれば logrotate コマンドを実行しなくても自動的にローテーションが実施されることを意味しているものではありません。
そのため、logrotate コマンドの実行タイミングによっては、指定サイズを大幅に超えるログファイルが作成される場合がありますので、注意してください。

5.2.3 システムログファイルのバックアップ

過去のシステム稼働状況や障害発生状況を調査する必要がある場合に備えて、システムログファイルを定期的にバックアップすることを推奨します。
システムログファイルのバックアップ処理は、システム運用の一環として計画／実施してください。
また、システムログファイルのバックアップにおいては、必ずローテーションが実施されたあとの過去世代ファイル（切り替え済みシステムログファイル）をバックアップ対象としてください。
稼働中のシステムログファイルをバックアップ対象とすると、バックアップ中にシステムログファイルが更新されたり、バックアップ時点では過去世代ファイルに移行してしまったログがバックアップされなかったりなど、連続したログの蓄積とならない問題が生じます。
このことから、システムログファイルのバックアップにおいては、ログローテーション後の世代ファイルをバックアップ対象とするよう、運用の設計を行ってください。

参 考

- **システムログファイルのローテーションにおいては、過去世代ファイルは以下のように管理されます。システムログファイルの世代ファイルのバックアップ運用時の参考としてください。**
  また、**「5.3.2 ユーザー固有ログファイルのローテーション運用方式」（→ P.98）****も併せて参照してください。**
  - **設定ファイルに指定された保有世代数分の世代ファイルが、稼働中システムログファイルと同じディレクトリ配下に作成されます。**
  - **世代ファイルはファイル名で識別され、システムログファイル名 + ". 番号 " というファイル名となります。**
  - **世代ファイルに付与される番号は、常に世代の若いファイルから順番に、1 から番号が付与されます。これは言い換えると以下を意味します。**
    - **世代ファイルは、ログローテーションが実施されるたびに、付与された番号が 1 つずつ増加しファイル名が変わる。**
    - **ログローテーションが実施され、稼働中から非稼働に移行したシステムログファイルは、番号に 1 番が付与される。**
  - **ログローテーションによる世代ファイルの番号付けの結果、保有世代数を超えた番号のファイルは自動削除されます。**

5.3 ユーザー固有ログファイルのローテーション運用

ユーザー固有ログファイルに対してローテーション運用を実施する場合に、考慮する事項について説明します。

- **「5.3.1 ローテーション運用の検討事項」（→ P.97）**
- **「5.3.2 ユーザー固有ログファイルのローテーション運用方式」（→ P.98）**

参考

- **ユーザー固有ログファイルのローテーション運用の設定手順については、「5.5 ユーザー固有ログのローテーション運用設定」（→ P.104）を参照してください。**

5.3.1 ローテーション運用の検討事項

ユーザー固有ログファイルに対してローテーション運用を実施する場合は、以下の点を検討します。

- 対象ログファイル
  ローテーション運用を実施する対象のユーザー固有ログファイルを決定します。
- 実行ユーザー
  ユーザー固有ログファイルのパーミッションに基づき、どのユーザー ID で実施すべきかを決定してください。
  - 実行ユーザーが root 権限の場合
    デフォルトのシステムログファイルのローテーション運用に追加して行う方法があります。この場合の設定については、「5.5.1 システムログのローテーション運用に追加する場合」（→ P.104）を参照してください。
  - 実行ユーザーが root 権限ではない場合、または root 権限においても独自でローテーション運用を行いたい場合
    利用するコマンドおよびスケジューリングシステムの仕様に基づき、運用の設計を行ってください。
    logrotate コマンドを独自に運用する場合、および cron でスケジューリングを実施する場合の設定方法については、「5.5.2 システムログのローテーション運用とは別に行う場合」（→ P.105）を参照してください。

- ローテーション運用実行コマンド
  ローテーション運用を実行するコマンドを決定してください。システム標準コマンドは logrotate コマンドです。
  logrotate コマンドで実行する場合は、ログファイルを出力するアプリケーションの機能を調査し、ログファイルのローテーション運用をする際のファイル切り替え方式を決定してください。
  切り替え方式について詳しくは、「5.3.2 ユーザー固有ログファイルのローテーション運用方式」（→ P.98）を参照してください。

5.3.2 ユーザー固有ログファイルのローテーション運用方式

ユーザー固有ログファイルのローテーション運用実行コマンドとして logrotate コマンドを使用する場合、ローテーション運用方式を決定します。
ローテーション運用方式は利用者が設定ファイルに指定するため、事前に決定しておく必要があります。

■ ローテーション運用方式

ローテーション運用方式には、ユーザー固有ログファイル自体を切り替えて保存する方式と、ユーザー固有ログファイルのコピーを保存する方式があります。アプリケーションにログファイルのリロード機能があるかどうかにより、選択します。

参 考

- **ログファイルのリロード機能とは、アプリケーションが稼働中のままログファイルをいったんクローズし、再オープンする動作を、何らかの方法（シグナルの通知など）により外部から任意のタイミングで実行する機能です。**

- アプリケーションに、ログファイルのリロード機能がある場合
  ユーザー固有ログファイル自体を切り替えて保存する方式を使用します。ユーザー固有ログファイル自体を切り替えて使用する方式については、「■ ユーザー固有ログファイル自体を切り替えて保存する方式」（→ P.99）を参照してください。
- アプリケーションに、ログファイルのリロード機能がない場合
  ユーザー固有ログファイルのコピーを保存する方式を使用します。ユーザー固有ログファイルのコピーを保存する方式については、「■ ユーザー固有ログファイルのコピーを保存する方式」（→ P.100）を参照してください。

■ ユーザー固有ログファイル自体を切り替えて保存する方式

ログファイルを出力するアプリケーションが、ログファイルのリロード機能を有している場合に実施できる方式です。
指定方法は、logrotate 設定ファイルに「create」ディレクティブを記述します。
ユーザー固有ログファイルのローテーション運用は、以下のとおり実現されます。

1 **世代ファイルに付与した番号をカウントアップして、ファイル名を変更**

2 **現在運用中のユーザー固有ログファイル名を、一番若い番号を世代ファイル名に変更**

3 **ユーザー固有ログファイル名で新規ファイルを 0 バイトサイズで作成**

4 **アプリケーション／システムに、ユーザー固有ログファイルを新規作成したファイルをリロードするように通知**

5 **世代管理から外れたユーザー固有ログファイル（番号の一番大きいユーザー固有ログファイル）を削除**

例えば、「app.log」というログファイルに対し、2 世代分の旧ログファイルを保持するよう設定した場合、「app.log.1」および 「app.log.2」というファイルが、ユーザー固有ログファイルのローテーション運用により自動的に作成されます。
このときのローテーション運用の動作は、以下のとおりです。

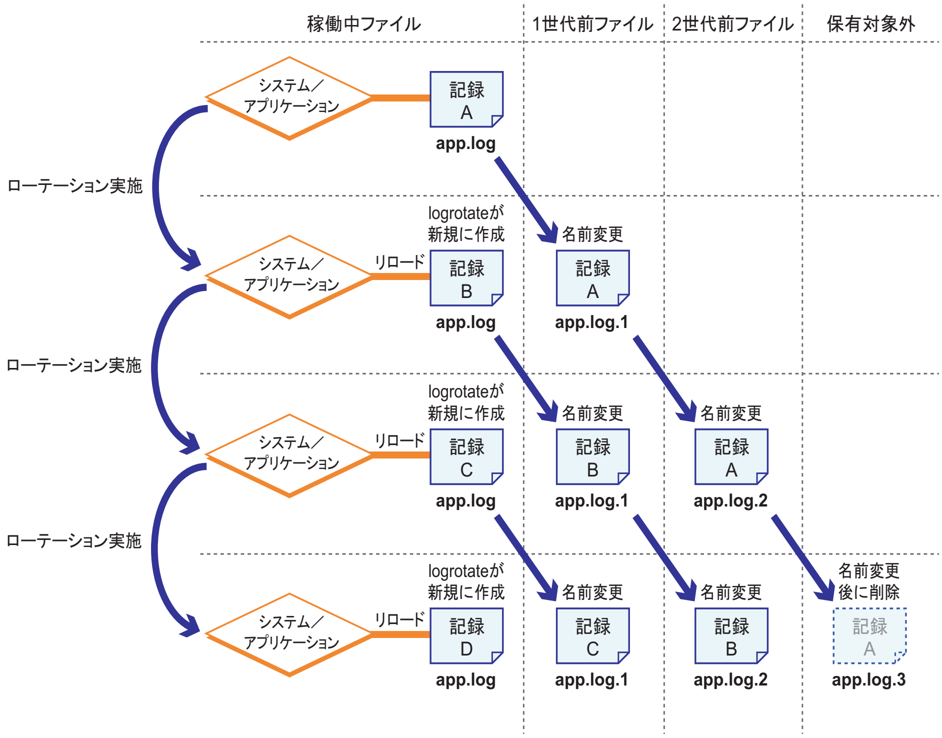


図 5 - 3：ローテーション運用の動作

図 5 - 3 のように、世代管理を超えた最も古いユーザー固有ログファイルは、ログローテーション運用のたびに自動的に削除されます。よって、ユーザー固有ログファイルをすべて保持しておくためには、次回のローテーションまでの間に、世代管理における最も古いユーザー固有ログファイルのバックアップが完了している必要があります。運用においては、ローテーション運用が実施された段階で、切り替え直後の世代ファイル（図 5 - 3 では app.log.1）をバックアップすることを推奨します。

■ ユーザー固有ログファイルのコピーを保存する方式

ユーザー固有ログファイルのコピーを保存する方式は、ログファイルを出力するアプリケーションが、アプリケーションを終了するまでログファイルを常にオープンしたままの状態で運用する仕様の場合に実施します。指定方法は、logrotate 設定ファイルに「copytruncate」ディレクティブを記述します。
ユーザー固有ログファイルのコピーによるローテーション運用は、以下のとおり実現されます。

1. **世代ファイルに付与した番号をカウントアップして、ファイル名を変更**
2. **一番若い番号の世代ファイル名で、新規にファイルをオープン**
3. **ユーザー固有ログファイルの内容を読み込み、手順 2 でオープンしたファイルにコピー**
4. **ユーザー固有ログファイルを 0 サイズにクリア**
5. **世代管理から外れたユーザー固有ログファイル（番号の一番大きいユーザー固有ログファイル）を削除**

このときのローテーション運用の動作は、以下のとおりです。

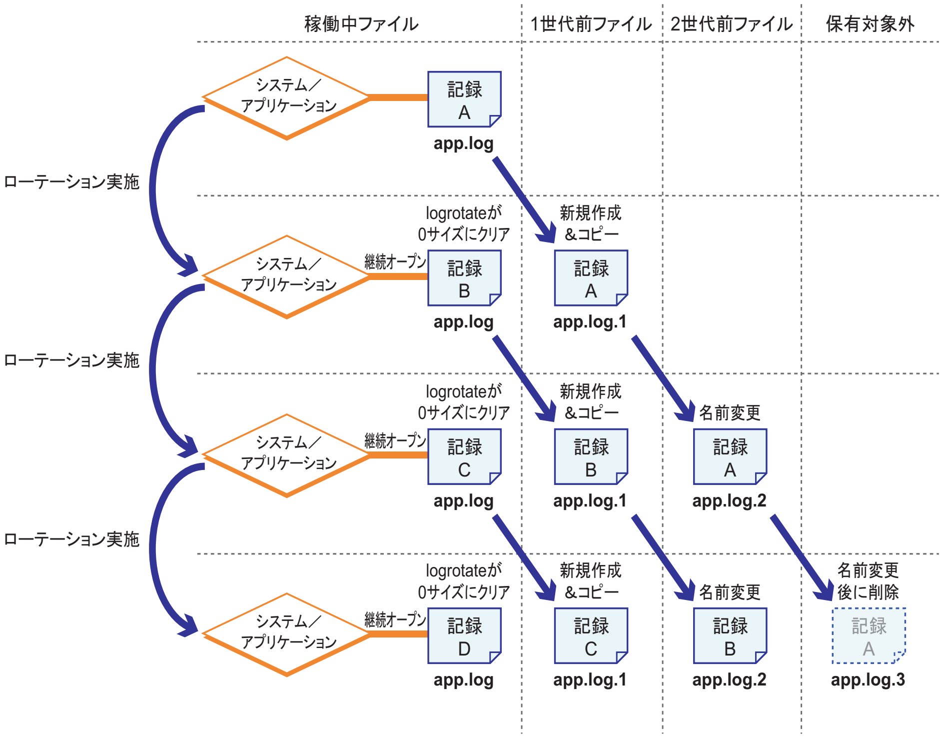


図 5 - 4：ローテーション運用の動作

世代管理を超えた最も古いユーザー固有ログファイルは、ローテーション運用のたびに自動的に削除されます。
よって、ユーザー固有ログファイルをすべて保持しておくには、次回のローテーションまでの間に、世代管理における最も古いユーザー固有ログファイルのバックアップが完了している必要があります。運用においては、ローテーション運用が実施された段階で、切り替え直後の世代ファイル（図 5 - 4 では app.log.1）をバックアップすることを推奨します。

注 意

- **この切り替え方式では、ユーザー固有ログファイルのコピー終了後、0 サイズにクリアするまでのわずかの間にアプリケーションによるログ出力が行われた場合、そのログ出力が失われる可能性があります。**
  この方式を採用する前に、業務運用においてアプリケーションの停止／再起動が可能であるか否かを検討してください。アプリケーションの停止／再起動が可能な場合は、ユーザー固有ログファイル自体の切り替え方式による運用を行ってください。
  また、やむを得ずこの切り替え方式を採用する場合には、切り替え時にログ出力が失われる可能性を認識のうえ、ログ出力が行われない時間帯を選んでローテーション運用を行うなどの対応を検討してください。
- **この切り替え方式では、ユーザー固有ログファイルを 0 サイズにクリアしたあと、ファイルポインターがファイルの先頭に戻されません。そのため、ログファイルを出力するアプリケーションは、ファイルポインターを先頭に移動させてからログを出力する必要があります。**

5.4 システムログのローテーション運用設定

システムログのローテーション運用を変更する際の各種設定方法について説明します。

- 「5.4.1 スケジュールの変更」(→ P.102)
- 「5.4.2 切り替え周期などの変更」(→ P.103)

5.4.1 スケジュールの変更

システムログのローテーション運用のスケジュールを変更するには、日次処理全体の時刻を変更する方法と、ローテーション運用を日次処理から外して、個別に実施時刻を設定する方法があります。

■ 日次処理全体を変更する場合

/etc/crontab ファイルの日次処理の設定時刻を変更します。

1 **/etc/crontab ファイルの編集を開始します。**

```
vi /etc/crontab
```

2 **日次処理の設定時刻を変更して保存し、編集を終了します。**

```
     :
02 4 * * * root run-parts /etc/cron.daily
     :
```

「02 4」の部分は、毎日午前 4 時 2 分に日次処理が実行されることを意味しています。実施したい時刻に変更します。変更後は cron により自動認識されます。

■ ローテーション運用実施を日次処理から外して運用する場合

日次処理からローテーション運用を実施しているスクリプトを外し、別スケジュールを設定します。

1 **root 権限で、/etc/cron.daily 配下にある logrotate スクリプトを、任意のディレクトリ配下へ移動します。**

ただし、/tmp 配下など、一定期間後にファイルが削除されるようなディレクトリへは移動しないでください。

例：/root/daily ディレクトリを作成して運用する場合

```
# mv /etc/cron.daily/logrotate  /root/daily
```

2 logrotate **スクリプトの実施スケジュールの編集を開始します。**

```
# crontab -e
```

環境変数 EDITOR または VISUAL に設定されている場合はそのエディタープログラムで、設定されていない場合は vi エディターで、logrotate スクリプトの実施スケジュールの編集が開始されます。

3 logrotate **スクリプトの実施スケジュールを定義します。**

例：毎日 0 時 30 分に実行する場合

```
     :
30 0 * * * /root/daily/logrotate
     :
```

4 **変更後の設定内容を確認します。**

```
# crontab -l
```

5.4.2 切り替え周期などの変更

切り替え周期や保有世代数などを変更する場合は、各ログファイルに対応した設定ファイルを編集します。
設定ファイルの記述については、「5.6.2 設定ファイルの記述様式」（→ P.108）を参照してください。

1 root **権限で、ログファイルに対応する** /etc/logrotate.d **配下の設定ファイルをエディターで編集します。**

例：named デーモンのログ設定ファイルを vi エディターで編集する場合

```
# vi /etc/logrotate.d/named
```

例：ローテーション運用を毎日行う設定を追加する場合

```
/var/log/named.log {
    daily
    missingok
    create 0644 named named
    postrotate
        /sbin/service named reload 2> /dev/null > /dev/null
    endscript
}
```

2 logrotate **コマンドをデバッグモードで手動起動し、設定ファイルの変更によりエラーが発生しないことを確認します。**

```
# /usr/sbin/logrotate -d /etc/logrotate.conf 2>&1
```

記述にエラーがある場合、「error: ～」の警告が表示されますので、エラー箇所を修正し、再度確認してください。

5.5 ユーザー固有ログのローテーション運用設定

ユーザー固有ログのローテーション運用の設定方法について説明します。

- 「5.5.1 システムログのローテーション運用に追加する場合」(→ P.104)
- 「5.5.2 システムログのローテーション運用とは別に行う場合」(→ P.105)

5.5.1 システムログのローテーション運用に追加する場合

実行ユーザーが root 権限の場合は、システムログのローテーション運用に、ユーザー固有ログのローテーション運用の設定を追加できます。

1 **ログローテーションを実施したいユーザー固有ログファイルに対して、対応する設定ファイルを作成します。**

設定ファイルの名称は、対応するログファイルが連想できるファイル名にしてください。
また、設定ファイルの記述においては、必ず以下のようにログファイルの指定を行い、個別設定を行ってください。
ユーザー固有ログファイルが /home/user/user.log の場合の記述例です。

```
/home/user/user.log {
    nocompress
    missingok
}
```

- 個別設定内容において、切り替え周期、保有世代数などを省略した場合には、/etc/logrotate.conf に記述されているグローバル定義が継承されます。
- 設定ファイルの記述については、「5.6.2 設定ファイルの記述様式」(→ P.108)を参照してください。

2 **作成した設定ファイルを /etc/logrotate.d 配下に配置します。**

ファイルの読み取り権限を、読み取り可能に設定してください。

3 **logrotate コマンドをデバッグモードで手動起動し、設定ファイルの変更によりエラーが生じないことを確認します。**

```
# /usr/sbin/logrotate -d /etc/logrotate.conf
```

- 「reading config file *<設定ファイル名>*」という表示が続きます。配置したファイルが表示されることを確認してください。
- 記述にエラーがある場合、「error: ～」の警告が出力されます。エラー箇所を修正してください。

5.5.2 システムログのローテーション運用とは別に行う場合

個別にユーザー固有ログファイルのローテーション運用を行う場合で、かつ logrotate コマンドを使用して実施する場合の設定方法について説明します。

参 考

状態ファイルとは

- **logrotate コマンドの設定ファイルで指定したログファイルに対して、ローテーション運用を実施した最終日付を記録、保持するファイルです。[-s] オプションで指定します。**
  [-s] オプションで指定したファイルが存在しない場合は、そのファイル名で自動作成されます。
  [-s] オプションで指定しない場合は、/var/lib/logrotate.status ファイルが使用されます。

1 **ローテーション運用を実施したいユーザー固有ログファイルに対して、対応する設定ファイルを作成します。**

設定ファイルは任意の場所に作成します。設定ファイルの記述方法については「5.6.2 設定ファイルの記述様式」（→ P.108）を参照してください。

2 **状態ファイルを作成するディレクトリと、状態ファイルのファイル名を決定します。**

実際にファイルを作成する必要はありません。logrotate コマンドが自動的に作成します。

注 意

- **設定ファイルの作成場所は、標準設定ファイル（/etc/logrotate.conf）内の #include 文で指定しているディレクトリ以外を指定してください。標準設定ファイルについては、「5.6.1 設定ファイルの種類」（→ P.107）を参照してください。**

3 **logrotate コマンドをデバッグモードで手動起動し、設定ファイルの変更によりエラーが生じないことを確認します。**

以下のコマンドを実行します。状態ファイルの指定を忘れずに行ってください。

```
# /usr/sbin/logrotate -d -s <状態ファイルフルパス名> <設定ファイルフルパス名>
```

・デバッグモードで実行した段階では、まだ状態ファイルは作成されません。
・記述にエラーがある場合、「error: 〜」の警告が出力されますので、エラー箇所を修正してください。

4 **logrotate コマンドを手動起動し、正常にローテーション運用が実施されることを確認します。**

以下のコマンドを実行します。状態ファイルの指定を忘れずに行ってください。

```
# /usr/sbin/logrotate -s <状態ファイルフルパス名> <設定ファイルフルパス名>
```

5　スケジュール運用を行う場合は、手順 4 で実行したコマンドを登録します。

cron でスケジュール運用を行う場合は、「■ ローテーション運用実施を日次処理から外して運用する場合」（→ P.102）の作業手順を参考に、crontab コマンドを使用して登録してください。

注 意

- **個別にユーザー固有ログファイルのローテーション運用を行う場合は、以下の点に注意してください。**
  - **設定ファイルを、システムのデフォルトとは別に用意する。**
    「5.6.1 設定ファイルの種類」（→ P.107）
  - **状態ファイルを、システムのデフォルトとは別に用意する。**
  - **logrotate コマンド実行時に、設定ファイルに加えて、状態ファイルの指定を行う。**

5.6 ローテーション運用の設定ファイル

logrotate コマンドを使用して、ローテーション運用を行う場合に必要な設定ファイルの種類と、記述様式について説明します。

- 「5.6.1 設定ファイルの種類」（→ P.107）
- 「5.6.2 設定ファイルの記述様式」（→ P.108）

5.6.1 設定ファイルの種類

logrotate コマンドは起動時に設定ファイルを読み込み、ローテーション運用の実施条件を取得します。logrotate コマンドが読み込む設定ファイルには、以下の 2 種類があります。

- 標準設定ファイル（/etc/logrotate.conf）
  logrotate コマンドが起動時にデフォルトで読み込む設定ファイル
- 個別設定ファイル（/etc/logrotate.d ディレクトリ配下のファイル群）
  - 標準設定ファイルとは別に、個別にログファイルの設定値を変更する設定ファイル
  - 「/etc/logrotate.d」は、標準設定ファイル内の include ディレクティブでデフォルト値として設定

図 5 - 5 に logrotate コマンドと各設定ファイルの関係を示します。

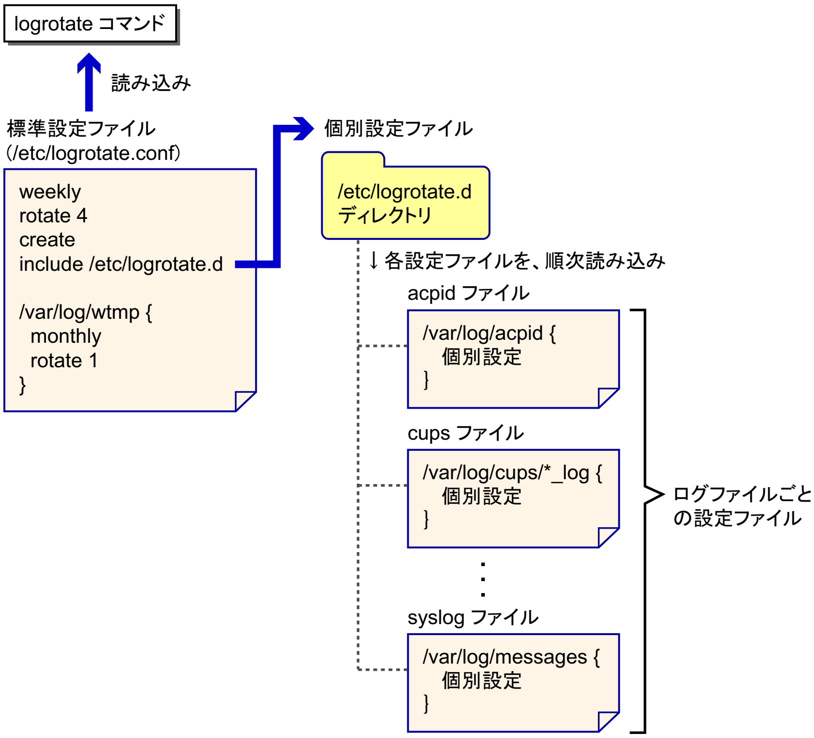


図 5 - 5：logrotate コマンドと各設定ファイルの関係

5.6.2 設定ファイルの記述様式

logrotate コマンドの設定ファイルには、ログファイルの切り替え、および世代管理の制御を行うディレクティブなどを記述します。

参 考

- **ディレクティブについては、logrotate コマンドの仕様（man コマンド）を参照してください。**

図 5 - 6 に設定ファイルの記述様式を示します。

logrotate コマンドの設定ファイル書式

```
[ディレクティブ1]
[ディレクティブ2]          } 各ログファイルに共通の設定
        :
ログファイル名 [ログファイル名 …] {
    [ディレクティブ①]
    [ディレクティブ②]      } ログファイルごとに個別の設定
            :
}
        :
```

図 5 - 6：logrotate コマンドの設定ファイル書式

設定ファイルには、以下の決まりがあります。

- 全ログファイルに共通の設定を先に記述し、そのあとにログファイルごとの個別設定を記述します。
- 共通設定で設定された項目を、個別設定によって上書きできます。
- ログファイル名には、複数のファイル名を併記したり、ワイルドカード（*）を使用したりできます。
  この場合、該当する複数ファイルに対して共通の設定が適用されます。
- 「#」で開始する行はコメント行とみなされ、行末まで無視されます。

標準設定ファイル（/etc/logrotate.conf）の記述例

```
weekly                       <---- 切り替え単位を週単位とする
rotate 4                     <---- 世代ファイルを4世代保持する
create                       <---- ログファイル自体を切り替えて保存する方式とする
include /etc/logrotate.d     <---- 個別設定ファイルの格納先ディレクトリを指定する

/var/log/wtmp {              <---- /var/log/wtmp用の個別設定の宣言
    monthly                  <---- 切り替え単位を月単位とする
    create 0644 root utmp    <---- ログファイル自体を切り替えて保存する方式とする
                                   (切り替え後に作成されるログファイルのアクセス権限などを指定)
    rotate 1                 <---- 世代ファイルを1世代保持する
}
```

図 5 - 7：標準設定ファイル（/etc/logrotate.conf）の記述例

個別設定ファイルの書式

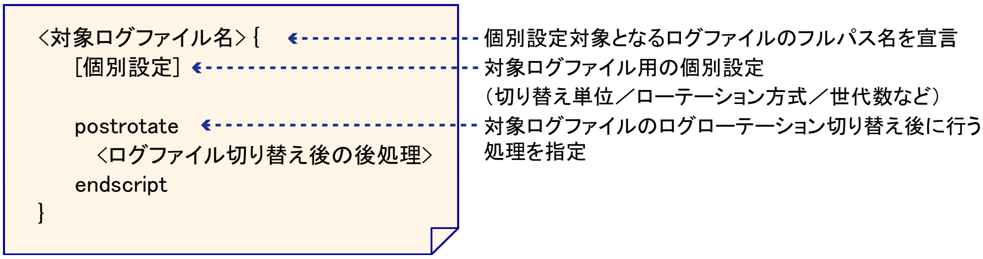


図 5 - 8：個別設定ファイルの書式

「postrotate」は、対象ログファイルのローテーション運用方式が「■ ユーザー固有ログファイル自体を切り替えて保存する方式」（→ P.99）の場合に指定してください。
「＜ログファイル切り替え後の後処理＞」は、対象ログファイルのクローズおよび再オープンを行う処理を記述してください。
以下は、対象ログファイルを操作するアプリケーションまたはサービス（以降、アプリケーションと表記します）側で、対象ログファイルのクローズおよび再オープンを行う場合の、postrotate の記述例です。ただし、アプリケーション側の処理は省略します。

- アプリケーションを停止できない場合
  アプリケーション内で HUP シグナルをハンドリングし、対象ログファイルのクローズおよび再オープンを行います。

```
postrotate
     kill -HUP `cat /home/user/user.pid`
endscript
```

  /home/user/user.pid には、アプリケーションのプロセス ID が設定されているものとします。

- アプリケーションを停止できる場合
  アプリケーションを再起動することで、対象ログファイルのクローズおよび再オープンを行います。

```
postrotate
     /home/user/userapply restart`
endscript
```

  /home/user/userapply はアプリケーションを表し、アプリケーションは [restart] オプションにより再起動するものとします。

注 意

- **設定ファイルを編集する場合は、以下の点に注意してください。**
  - **あるログファイルの定義が複数の個別設定ファイルに存在する場合、当該ログファイルのローテーション処理が失敗します。実際に利用しない個別設定ファイル（過去の設定の保存用ファイルなど）は削除するか、または個別設定ファイルの配置ディレクトリ以外に移動してください。**
  - **存在しないログファイルが設定ファイルに定義されていると、当該ログファイルのローテーション処理が失敗します。当該ログファイルが存在しない場合は、事前に空のログファイルを作成しておくか、または（実際のログ出力により）ログファイルが作成されたあとにローテーションを実施するようにしてください。**

6

第 6 章

システム設計：時刻補正の導入（NTP）

この章では、Linux **システムにおける時刻の仕組みや時刻補正の必要性、時刻補正に重要な役割を果たす** NTP（Network Time Protocol）**サービスについて説明します。**

6.1 時刻の仕組みと時刻補正の概要

Linux システムには、「ハードウェア時計」と「ソフトウェア時計」があります。
ここでは、それぞれの時計に対する時刻補正の概要について説明します。

- 「6.1.1 時刻補正の必要性」(→ P.111)
- 「6.1.2 ハードウェア時計とシステム時計」(→ P.111)

6.1.1 時刻補正の必要性

コンピュータは、その性能にかかわらず、内蔵されている水晶発振器によって時刻を計算しています。この水晶発振器は個体ごとに誤差があるため、時間の経過とともに正確な時刻からの誤差が累積し、正確な時刻が維持されません。この状態を放置した場合、次のような影響が生じます。

- イベントログの時刻ずれ
  イベントログの記録においてタイムスタンプが付与される場合、正確な時刻が記録されません。
- システム間連携における障害発生
  時刻が一致しないシステム間で、タイムスタンプを伴うデータの連携を行っている場合、時刻のずれによりシステムがエラーと判定する事象が発生する場合があります(ネットワーク共有しているファイルの更新時刻が、過去の時刻で更新されるように見えるなど)。

このことから、システムの時刻を正確に保つための時刻補正の仕組みが必要となってきます。

6.1.2 ハードウェア時計とシステム時計

ハードウェア時計は、ハードウェアに搭載されている時計です。リアルタイムクロックとも呼ばれています。システム停止時も、バッテリーにより駆動され、時刻を保持します。
システム時計は、システム稼働中にタイマー割り込みを使用してカーネルが時刻を管理しているソフトウェア時計です。OS は、システム時計を用いて時刻情報を返します。

Linux システムにおいては、ハードウェア時計とシステム時計は、以下のように関係します。

- システム起動時
  ハードウェア時計から時刻を取得し、システム時計に設定します。
- システム停止時
  システム時計の時刻をハードウェア時計に反映、更新します。

■ ハードウェア時計の補正

システムを長期稼働する（システムの起動／停止の繰り返しがない）場合は、ハードウェア時計とシステム時計の同期をとるタイミングがないため、ハードウェア時計はシステム時計との時刻のずれが累積していきます。これは、ハードウェアと OS の各イベントログのタイムスタンプに影響します。例えば、トラブル発生時の原因究明において、OS とハードウェアのそれぞれで発生したイベントの時系列上の比較が困難になります。そのため、システムを長期稼働する場合は、システム時計の時刻を用いて定期的にハードウェア時計の時刻を更新することを推奨します。

● ハードウェア時計の更新

root 権限で以下のコマンドを実行して、システム時計に合わせてハードウェア時計の時刻を更新します。

```
# /sbin/hwclock --systohc
```

● システム時計とハードウェア時計の時間差の確認

ハードウェア時計とシステム時計の時間差を確認するには、root 権限で以下のコマンドを実行します。

```
# /sbin/hwclock --show
<曜日><月><日><時刻><年><システム時計との時刻差> seconds ← 出力結果
```

以下は、出力の一例です。

```
Tue Oct 19 08:44:24 2010 -0.591105 seconds
```

表示される時刻は、ハードウェア時計の時刻です。そのとなりに表示されている秒数は、システム時計との差を示しています。マイナス値は、ハードウェア時計が進んでいることを意味します。時間差が大きい場合、「● ハードウェア時計の更新」（→ P.112）を実行してください。

■ システム時計の補正

システム時計も、正確な時刻を保つために時刻補正の仕組みが必要となります。本書では、NTP 運用による時刻補正を行うことを推奨します。NTP 運用については、「6.2 NTP 運用」（→ P.113）以降で詳しく説明します。

6.2 NTP 運用

RHEL では、時刻補正の仕組みとして NTP サービスを標準提供しています。NTP によるシステム時刻の補正の仕組みについて説明します。

- 「6.2.1 NTP とは」（→ P.113）
- 「6.2.2 NTP の時刻同期の仕組み」（→ P.114）
- 「6.2.3 RHEL5 における実装」（→ P.117）
- 「6.2.4 時刻補正の方法」（→ P.118）

6.2.1 NTP とは

Network Time Protocol（NTP）とは、ネットワークに接続される機器において、機器が持つ時計を正しい時刻へ同期するためのプロトコル（規格）です。
ネットワーク上に基準となる時刻情報を保有するサーバを設置し、そのサーバを基準として、クライアントの内部時計をそのサーバの時刻と一致させることにより、ネットワーク上のコンピュータに同じ時刻を設定します。
基準となる時刻には、協定世界時（UTC）があります。NTP を利用することにより、GPS などを使用して協定世界時へ同期しているサーバに、LAN ／ WAN 経由で、それぞれミリ秒／数 10 ミリ秒の精度で時刻を合わせることができます。
NTP は、OSI 基本参照モデルの第 7 層（アプリケーション層）に位置し、ネットワーク通信の待ち受けに UDP ポートの 123 番を使用します。

注 意

- **RHEL システム内外のファイアーウォール設定で UDP ポートの 123 番の通信が遮断されている場合は、NTP 運用に際しては遮断設定を解除してください。**

■ NTP の利点

コンピュータの時刻を正しい時刻に設定するために、date コマンドなどを用いて手動で時刻設定を行う場合、以下の問題が発生します。

- 人的作業が介在することにより、作業が煩雑となり、かつ頻繁な実施が困難
- 手作業に伴う誤差が発生

これに対して、自動的にコンピュータの時刻を補正する仕組みがあれば、上記の問題を解決できます。NTP は、この仕組みを実現するために設計されました。NTP 運用を行うことにより、手動設定より高い精度で時刻の補正を行うことができます。

6.2.2　NTP の時刻同期の仕組み

ここでは、NTP の時刻同期の仕組みについて、NTP サーバの階層構造と処理の流れを説明します。

■ 階層構造（stratum）

NTP のネットワークは、stratum と呼ばれる階層構造になっており、インターネット上でもイントラネット上でも stratum（階層）の番号がどれだけ小さいかで、どれだけ正確な時計（UTC）に近いかが表されます。最上位のサーバが正確な時計から標準時を取得し、下位のサーバはそれを参照することで時刻を合わせます。最上位の NTP サーバは stratum 1 で、階層を下りるごとに数字が 1 つずつ大きくなります。また、最下位の stratum は 16 となり、stratum 16 に同期することはできません。

図 6 - 1 に、時刻同期を行う NTP サーバの stratum 階層の概念図を示します。階層構造は、各サーバ機の業務上の関係性を考慮して構成します。例えば、ファイル共用サービスにおいて、ファイルサーバの時刻に各ワークステーションが NTP による時刻同期をする運用とした場合、ファイルサーバの階層が stratum 3 ならば、各ワークステーションの階層は stratum 4 という階層化がなされます。

最下位は stratum 16 なので、15 階層まで NTP サーバが存在できます。この参照関係を調整することで負荷分散が可能です。また、一般に stratum の大きさよりも、サーバとのネットワーク的な近さ（通過するルータの数など）のほうが時刻精度に大きく影響します。

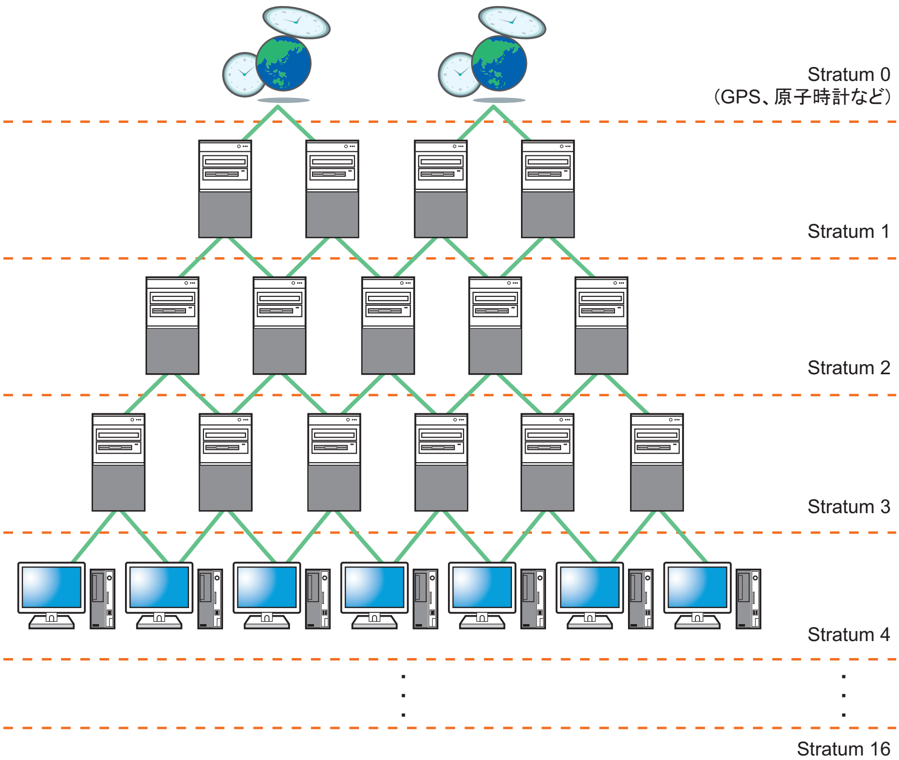


図 6 - 1：NTP の階層構造

■ 時刻補正処理の流れ

NTP の時刻補正の流れは、図 6 - 2 のようになります。
NTP サーバ機能が NTP クライアント機能を使って上位の NTP サーバに接続すると、上位 NTP サーバとのネットワーク通信の遅延を継続的に計測し、受け取った時刻情報を補正して自動的にミリ秒単位の精度で時刻を同期します。
NTP サーバ機能は時刻を上位サーバに合わせるだけでなく、自サーバより下位の NTP サーバから時刻を参照される側になることも可能です。その場合、stratum の値は上位サーバの stratum 値より 1 大きくなります。

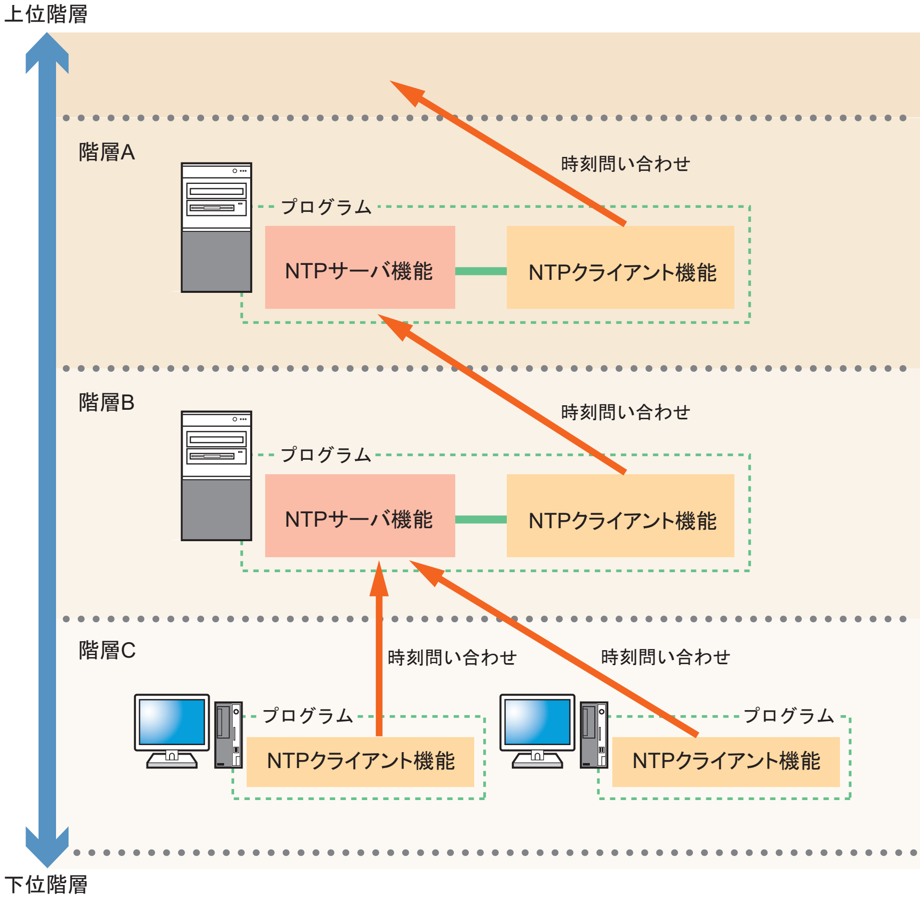


図 6 - 2：時刻補正処理の流れ

■ ローカルクロック

外部ネットワークと隔絶されたネットワーク環境で、NTP を用いてネットワーク内の機器の時刻を同期させたい場合は、時刻の基準となるシステムを用意する必要があります。このとき時刻の基準となるシステムは、NTP サービスを稼動させたうえで、NTP の時刻を自システムのカーネルの時計に同期させるという運用形態を取ります。この自システムのカーネル時計のことを、NTP 運用ではローカルクロックと呼びます。
また、NTP 運用では、外部の NTP サーバとローカルクロックを併用して指定することができます。この場合、通常は stratum 値の小さい外部 NTP サーバと同期し、ローカルクロックと同期することはありません。しかし、NTP サーバが起動していない場合や、ネットワークが不安定などの理由で応答がない場合は、ローカルクロックとの同期に切り替わります。このとき、NTP サーバの stratum 値はローカルクロックに設定した stratum 値より 1 大きくなります。ローカルクロックの指定方法、および stratum 値の設定については、「6.4.1 /etc/ntp.conf」（→ P.130）を参照してください。

■ peer 設定

NTP サーバは、上位の Stratum に対して同期を行う設定に加え、同じ Stratum 階層内の NTP サーバ間において、互いを時刻同期の対象として設定することができます。この方法を peer 設定と呼びます。peer 設定を行う際は、以下の条件を満たす必要があります。

- 同一 Stratum 階層内の NTP サーバであること
- peer 設定によってペアを組む同一階層のサーバ間で、同一の上位 NTP サーバを参照してはならない

peer 設定を行うことで、以下の効果を得ることができます。

- peer 設定をした NTP サーバ間で、時刻の差が収束されます。これにより、同一 Stratum 階層内の NTP サーバの時刻精度が均一化されます。
- peer 設定をした NTP サーバのいずれかが上位 NTP サーバと通信できなくなった場合も、peer 設定された NTP サーバと同期することにより、時刻精度が維持されます。

注 意

- **peer 設定の際の条件に従わない peer 設定（異なる Stratum 階層の NTP サーバ間や、同じ上位 NTP サーバを参照している NTP サーバ間の peer 設定）がなされた場合は、peer 設定された NTP サーバ間で時刻ずれが発生しますので、そのような設定は行わないでください。**

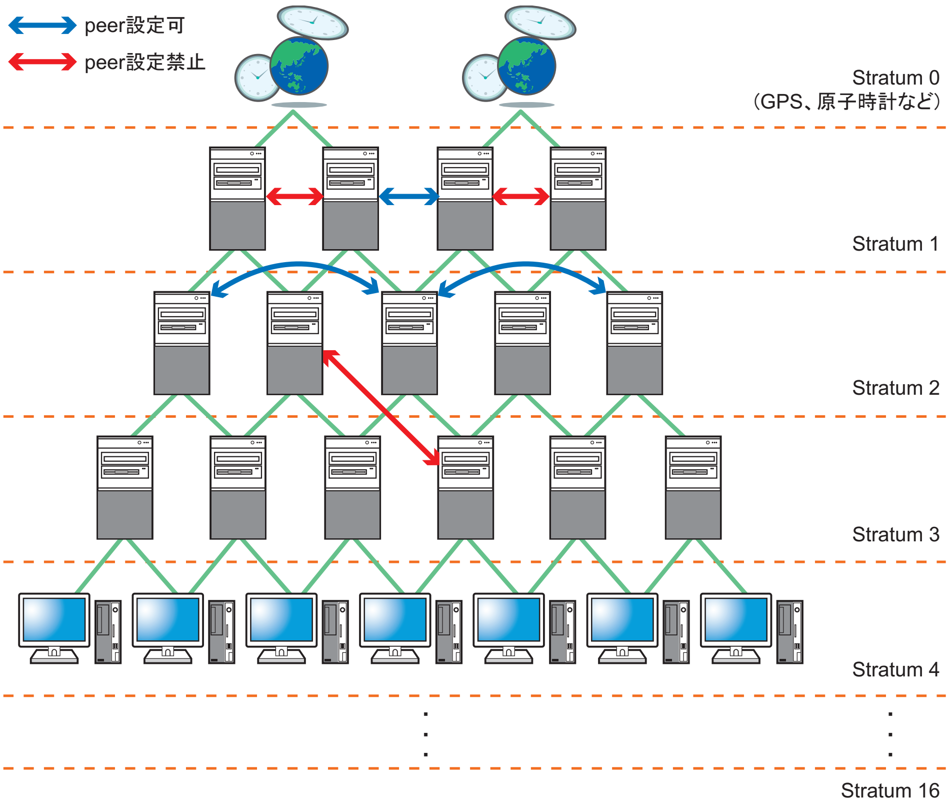


図 6 - 3：peer 設定が可能なサーバと禁止されるサーバ

peer 設定の方法については、「6.4.1 /etc/ntp.conf」（→ P.130）を参照してください。

6.2.3 RHEL5 における実装

RHEL には、NTP サーバ／クライアントとして ntpd、NTP クライアントとして ntpdate コマンドがあります。

■ ntpd

ntpd は、以下の 2 つの機能を持つデーモンです。

- NTP クライアントとして上位 NTP サーバと時刻同期、時刻補正を行うことができます。
- NTP サーバとして下位 stratum の NTP クライアントへ時刻サーバとして公開を行う機能を持っています。

ntpd では、複数の NTP サーバと接続できます。ntpd は、より多くの複数のサーバと時刻の比較やネットワークの遅延の計測を行い、重み付けを行って、最も信頼できると判断した時刻を現在時刻として利用します。

■ ntpdate コマンド

ntpdate コマンドは上位 NTP サーバへ時刻参照を行い、上位 NTP サーバの時刻と自身のシステム時刻を合わせます。ntpd も NTP クライアントとしての機能を持っていますが、ntpdate コマンドは常駐プログラムを減らしたい場合や、任意のタイミングで時刻補正を行いたい場合に有効です。

6.2.4　時刻補正の方法

■ 時刻問い合わせの動作

デフォルトの設定では、クライアントからは最初に 64 秒間隔で上位 NTP サーバに問い合わせを行います。この問い合わせ間隔をポーリング間隔といいます。時刻同期が完了すると、ネットワークのオーバーヘッドを減少させるために、ポーリング間隔を 128 秒、256 秒、512 秒、1024 秒と広げていき、最大 1024 秒まで延ばします。

■ Step 調整と Slew 調整

NTP の時刻補正には、Step 調整と Slew 調整の 2 つの方法があります。

● Step 調整

時刻のずれを一気に補正します。そのため、時刻が急激に変化し、時刻が逆進する可能性があります。

● Slew 調整

時刻のずれをゆっくりと（1 秒間に最大 0.5 ミリ秒ずつ）補正していきます。
NTP は、現時刻に定期的に加算する増分を調整することで時刻補正を行います。上位 NTP サーバと時刻を比較し、以下の調整を行います。

- NTP クライアント側が遅れている場合：増分を大きくする
- NTP クライアント側が進んでいる場合：増分を小さくする
- NTP サーバと時刻が等しい場合：増分は変更なし

このため、大幅な時刻のずれが補正されるまでに時間がかかります。

■ 時刻補正モード

ntpd を用いて時刻補正を行う場合、Step 調整と Slew 調整を組み合わせて使用します。また、ntpd 起動時の [-x] オプションの指定により、Step 調整と Slew 調整のタイミングを変えることができます。本マニュアルでは、[-x] オプションを使用せずに NTP 運用する場合を「Step モード」、[-x] オプションを使用して NTP 運用する場合を「Slew モード」と記載します。

● Step モード

Step モードは、ntpd デーモンのデフォルトの動作モードです。
上位 NTP サーバとの時刻差が 128 ミリ秒以内であれば Slew 調整を行い、128 ミリ秒を超える場合には Step 調整を行います（図 6 - 4）。具体的な調整は、以下のように行われます。

- 上位 NTP サーバとの時刻差が 128 ミリ秒以内の場合
  Slew 調整による時刻補正を行います。
- 上位 NTP サーバとの時刻差が 128 ミリ秒を越えた場合
  上位 NTP サーバの時間を 900 秒監視し、時刻差を継続して確認します。
  ± 128 ミリ秒以上の時刻のずれが解消されない場合、いったん NTP サーバとの同期を外し、Step 調整による時刻補正を行います。その後は最低 5 回の時刻問い合わせにより同期状態となります（同期にかかる時間はポーリング間隔やその他の条件により異なります）。

● Slew モード

Slew モードにおいても、Step 調整と Slew 調整を使い分けて時刻補正を行います。ただし、Step モードと比較して、Slew 調整による補正の範囲を拡大して時刻補正を行います。
上位 NTP サーバとの時刻差が 600 秒以内であれば Slew 調整を行い、600 秒を超える場合には Step 調整を行います（図 6 - 5）。これにより、時刻補正の大半は Slew 調整を用いて行われるようになります。
ntpd を Slew モードで運用する場合は、ntpd の起動オプションとして [-x] オプションを指定します。

ntpd の起動オプションの設定方法および確認方法については、「6.4.2 /etc/sysconfig/ntpd」（→P.136）を参照してください。

注 意

- **ntpd は、参照する NTP サーバとの時刻差が 1000 秒を越えると、Step モード／ Slew モードにかかわらず、以下のメッセージをシステムログファイル（/var/log/messages ファイル）に出力して終了します。**

```
time correction of <秒数> seconds exceeds sanity limit (1000); set clock
manually to the correct UTC time.
```

ただし、システム起動時に NTP サーバとの時差が 1000 秒以上であった場合には、Step モード／ Slew モードのそれぞれで以下の動作を行うため、ntpd 起動後すぐに終了することはありません。
- **Slew モード**
  NTP サービス開始時に ntpdate コマンドが実行され、時刻の step 補正が行われたあと、ntpd が起動します。
- **Step モード**
  ntpd 起動後の最初の時刻同期における step 補正では、1000 秒以上の時差があっても終了せずに補正を行います。

Step モード、Slew モードによる動作の違いは表 6 - 1 のとおりです。

表 6 - 1：起動モードによる ntpd の動作の違い

	Step モード（デフォルト）	Slew モード
ntpd 起動時	ntpdate コマンドを使用して、/etc/ntp/step-tickers ファイルに記述された NTP サーバの時刻に合わせます。	ntpdate コマンドを使用して、/etc/ntp/step-tickers ファイルに記述された NTP サーバの時刻に合わせます。 ただし、/etc/ntp/step-tickers ファイルに NTP サーバが記載されていない、またはファイル自体が存在しない場合には、/etc/ntp.conf の server ディレクティブに記載されている NTP サーバの時刻に合わせます。
ntpd 動作中	参照する NTP サーバとの時刻差が 128 ミリ秒以内の場合、1 秒間に最大 0.5 ミリ秒ずつ時刻補正を行います。参照する NTP サーバとの時刻差が 128 ミリ秒を超えた場合に、1000 秒までの時刻補正を一度に行います。1000 秒を超えた場合、システムログにメッセージを出力して終了します。	参照する NTP サーバとの時刻差が 600 秒以内の場合、1 秒間に最大 0.5 ミリ秒ずつ時刻補正を行います。参照する NTP サーバとの時刻差が 600 秒を超えた場合に、1000 秒までの時刻補正を一度に行います。1000 秒を超えた場合、システムログにメッセージを出力して終了します。
利点	一気に時刻補正を行うため、正しい時刻を即設定することができます。	時刻の進み具合を調整することで時刻補正を行うため、時刻は逆進しません。
欠点	時刻が逆進する場合があるため、データベースや時刻の逆進に対応していないアプリケーションなどに問題が発生する可能性があります。	1 秒間に最大 0.5 ミリ秒までしか補正できないため、時刻の調整に時間がかかる場合があります。

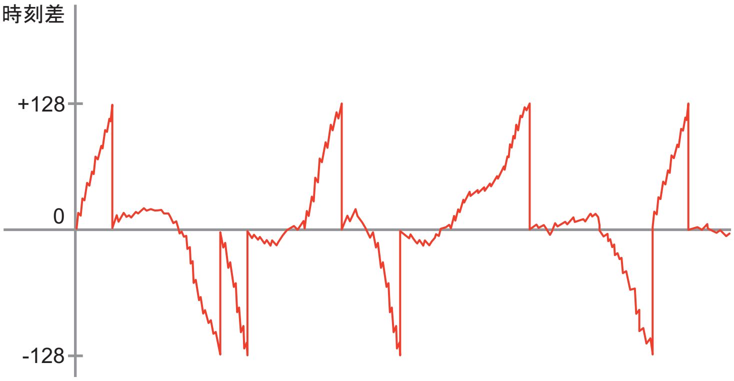


図 6 - 4：Step モードでの時刻補正

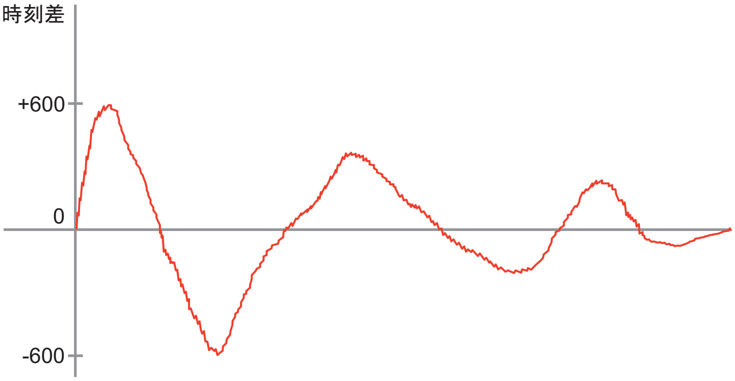


図 6 - 5**：**Slew **モードでの時刻補正**

6.3 NTP 環境の設計・導入

ここでは、ntpd を用いた NTP 環境の構築について、設計から導入までの流れを説明します。

- 「6.3.1 NTP 環境の設計手順」（→ P.122）
- 「6.3.2 参照する上位 NTP サーバの構成・参照設計」（→ P.123）
- 「6.3.3 時刻補正モードの選定」（→ P.125）
- 「6.3.4 起動時の初期参照 NTP サーバの設計」（→ P.126）
- 「6.3.5 高度な動作設計」（→ P.126）
- 「6.3.6 NTP 環境の導入」（→ P.127）

6.3.1 NTP 環境の設計手順

ntpd は、NTP サーバおよび NTP クライアントの機能を有しています。NTP サーバとして使用する場合には、正確な時刻情報を保有していることが求められます。通常はサイト内に代表 NTP サーバを置き、他の NTP サーバがそのサーバから時刻情報を得るのが一般的です。
NTP により時刻同期を行う場合には、次の流れで NTP の設定を行う必要があります。

1. 「6.3.2 参照する上位 NTP サーバの構成・参照設計」（→ P.123）

↓

2. 「6.3.3 時刻補正モードの選定」（→ P.125）

↓

3. 「6.3.4 起動時の初期参照 NTP サーバの設計」（→ P.126）

↓

4. 「6.3.5 高度な動作設計」（→ P.126）
ntpd 起動オプション、アクセス制御など必要に応じて実施します。

↓

5. 「6.3.6 NTP 環境の導入」（→ P.127）

6.3.2 参照する上位 NTP サーバの構成・参照設計

時刻同期の対象となる NTP サーバの構成については、必要となる要件があります。

■ 参照 NTP サーバの複数台指定

安定した NTP 運用を行うために、各 NTP クライアントから参照する外部の NTP サーバを複数台（3 台以上）指定してください。
ntpd は複数の NTP サーバを参照する際、その中からネットワークの距離や応答遅延時間、stratum などを考慮して似た誤差特性の NTP サーバを選び、その中で一番誤差が少ない NTP サーバと同期します。この似た誤差特性の NTP サーバを選ぶアルゴリズムをクラスタリング・アルゴリズムと呼びます。これにより、同期中の NTP サーバが同期に失敗したとしても、すぐにクラスタの中から代わりの NTP サーバを選ぶことができ、安定した同期状態を保つことができます。クラスタに属する NTP サーバを 3 台以上指定することで、1 台の時刻情報が異常な値を示しても、その NTP サーバを容易に排除できるようになります。
一方、NTP サーバの指定が 2 台以下である場合、時刻の正否が判断できなくなる場合があります。以下に、その一例を示します。

- NTP サーバが 1 台の場合
  NTP サーバとの接続が切断された場合に、同期対象の NTP サーバがなくなるため、NTP クライアントは時刻の正確さが保てなくなります。
- NTP サーバが 2 台の場合
  2 台の NTP サーバの時刻差が大きくなった場合に、NTP クライアントはどちらの NTP サーバの時刻がより正確であるかの判断ができなくなり、NTP クライアントの時刻が不安定になる場合があります。

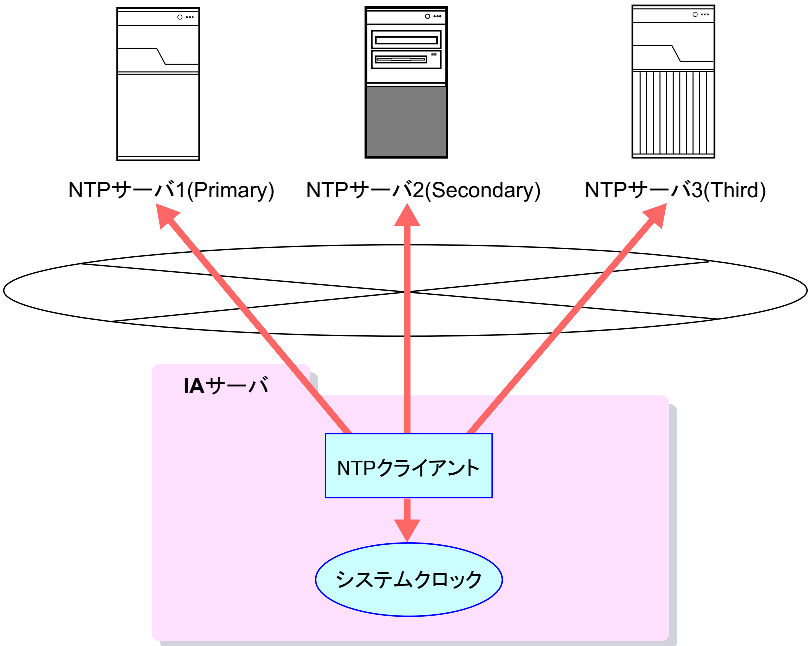


図 6 - 6：NTP サーバを複数台指定する場合の概念図

NTP サーバの指定は、/etc/ntp.conf 設定ファイルに設定します。設定方法については、「6.4.1 /etc/ntp.conf」（→ P.130）を参照してください。

■ PRIMEQUEST 1000 シリーズにおける特別な考慮

PRIMEQUEST 1000 シリーズでは、筐体内に装備されている MMB も NTP クライアント機能を有しており、外部の NTP サーバを用いて時刻補正を行うことができます。MMB が NTP 運用による時刻補正を行っている場合は、MMB とパーティションの時刻の差を小さくするために、MMB が指定する NTP サーバを各パーティションでも指定してください。

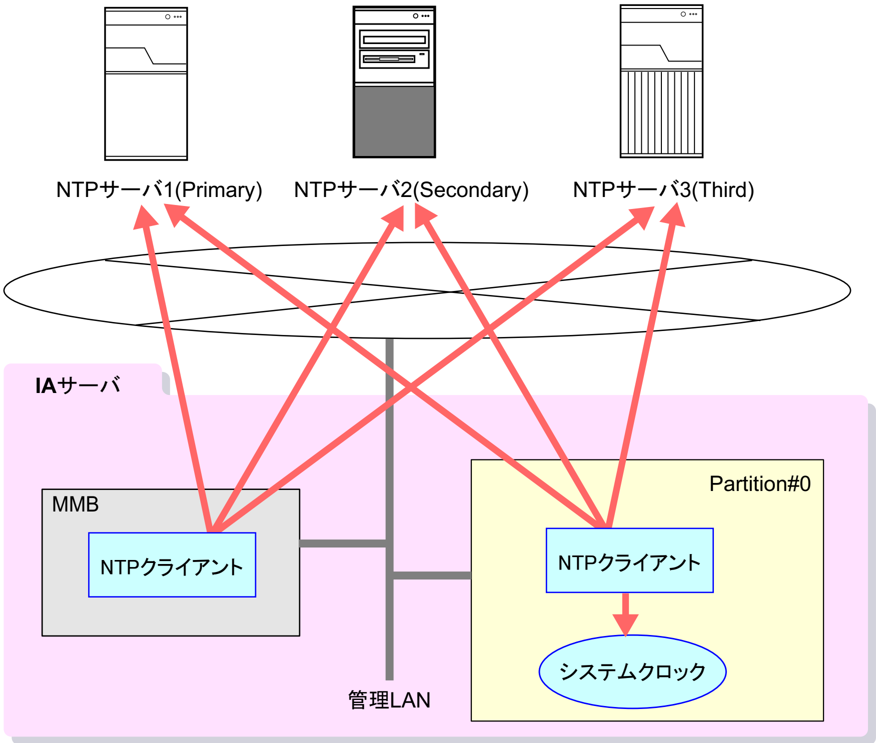


図 6 - 7：NTP サーバを利用した場合の概要図

■ サイト内代表 NTP サーバの複数台設置

ntpd は、NTP クライアント機能に加え、NTP サーバ機能を有しています。NTP サーバとして使用する場合には、正確な時刻情報を保有していることが求められます。通常は、サイト内に外部 NTP サーバと同期する代表 NTP サーバを複数台設置し、他の NTP サーバがそのサーバから時刻情報を得るのが一般的です。

6.3.3 時刻補正モードの選定

ntpd を用いて時刻補正を行う場合、Slew モードまたは Step モードのどちらの時刻補正モードを用いるかを、あらかじめ検討する必要があります。時刻補正モードについては、「■ 時刻補正モード」（→ P.118）を参照してください。デフォルトでは、Step モードが選択されます。
時刻補正モードの選択は、そのシステムで行われる業務の性格に影響を受けます。以下の点を選択基準の例として検討してください。

- 運用中の時刻の逆進が許容されない場合、Slew モードの選択を検討
- 正しい時刻からのズレを最小にしたい場合、Step モードの選択を検討

Slew モードを選択した場合、ntpd の起動時に [-x] オプションを指定する必要があります。起動オプションの設定方法については、「6.4.2 /etc/sysconfig/ntpd」（→ P.136）を参照してください。

■ Slew モード運用時の注意事項

● 上位 NTP サーバに関する注意事項

NTP サーバを Slew モードで運用する場合には、Step モードで運用している上位 NTP サーバを参照しないように設定してください。上位 NTP サーバが Step モードの場合、さらに上位の NTP サーバとの時刻差が 128 ミリ秒を超えた時点で Step 調整を行うため、その NTP サーバを参照している Slew モードで運用している NTP クライアントとの時刻差が大きくなる可能性があります。

● ntpd の再起動に関する注意事項

Slew モードの場合、ntpd の再起動時（起動時を含む）には、まず NTP サーバの時刻を一気に補正する処理が実行されます（起動処理中に ntpdate コマンドが自動実行されます）。これにより、ntpd を再起動する際には、時刻の逆進が発生する可能性があります。
ミドルウェア、アプリケーション（主に DB ソフトやクラスタソフト）が稼働している状態で ntpd を再起動した場合、時刻の逆進が発生すると処理異常となり、システムへ多大な影響を及ぼす可能性があります。
もし、運用中のシステムでミドルウェアやアプリケーションが稼働している場合は、ntpd の再起動を実施する前にそれらのサービスを停止し、ntpd を再起動したあとにサービスを起動してください。

6.3.4 起動時の初期参照 NTP サーバの設計

ntpd の運用においては、/etc/ntp/step-tickers ファイルが存在した場合、ntpd の起動時に ntpdate コマンドを実行します。
/etc/ntp/step-tickers ファイルが存在しない場合や、ファイル内に NTP サーバの記述がない場合は、Step モードでは ntpdate コマンドを実行しません。一方、Slew モードでは、/etc/ntp.conf に指定された NTP サーバを参照して ntpdate コマンドを実行します。詳細については、「6.4.3 /etc/ntp/step-tickers」（→ P.138）を参照してください。
Step モードにおいて、ntpd の起動時に ntpdate コマンドで時刻補正を行いたい場合には、/etc/ntp/step-tickers ファイルを作成して、/etc/ntp.conf ファイルで指定した NTP サーバを記述してください。

6.3.5 高度な動作設計

ntpd による時刻補正は、参照 NTP サーバの指定と時刻補正モードが必須の設計項目です。さらに高度な動作を行う場合は、必要に応じて動作設計を実施してください。

表 6 - 2：高度な動作設計の例

設計内容	設定方法
・参照する NTP サーバの優先度設定 [注 1] ・ローカルクロックを優先する設定（stratum の設定） ・NTP クライアントのアクセス制御 ・peerstatus/loopstatus/drift ファイルの配置先変更	「6.4.1 /etc/ntp.conf」（→ P.130）を参照してください。
・ntpd 起動時の起動オプションの追加 ・ntpd 起動時に実行される ntpdate コマンドのオプション追加 ・ntpd 起動時のハードウェアクロックとの時刻同期	「6.4.2 /etc/sysconfig/ntpd」（→ P.136）を参照してください。
・drift 値の設定／解除	「6.4.4 /var/lib/ntp/drift」（→ P.138）を参照してください。

［注 1］：参照する NTP サーバの優先度を設定する場合、Stratum 値ではなく、参照サーバの設定において prefer 値を指定してください。Stratum 値とは、NTP サーバの階層を示すもので、Stratum1 が最上位の階層になります。NTP クライアントは Stratum 値の低い NTP サーバを優先して参照しますが、複数台の上位 NTP サーバの参照優先度を指定するために Stratum 値を用いることは、本来の利用方法ではありません。NTP サーバの参照優先度の設定については、「6.4.1 /etc/ntp.conf」（→ P.130）を参照してください。

6.3.6 NTP 環境の導入

これまでの NTP 環境の設計に基づき、実際の導入作業を行います。

1 NTP サービスを停止します。以下のコマンドを実行し、NTP サービスの稼動状況を確認します。

```
# /sbin/service ntpd status
```

以下の出力結果の場合、NTP サービスが稼働していません。手順 2 に進んでください。

```
ntpd is stopped
```

以下の出力結果の場合、NTP サービスが稼働中です。

```
ntpd (pid <PID>) is running...
```

この場合、以下のコマンドを実行し、稼動している NTP サービスを停止してください。

```
# /sbin/service ntpd stop
```

以下の出力結果の場合、NTP サービスは停止しています。

```
Shutting down ntpd:                             [  OK  ]
```

2 NTP 環境の設定ファイルを編集します。

これまでの設計に基づき、各種設定ファイルを編集してください。最低限の設定項目として、以下の設定を行ってください。

参照する上位 NTP サーバの設定

NTP コンフィグファイル（/etc/ntp.conf）に NTP サーバの IP アドレスを以下のように設定します。

```
server  <信頼できるNTPサーバのIPアドレス-1>
server  <信頼できるNTPサーバのIPアドレス-2>
server  <信頼できるNTPサーバのIPアドレス-3>
```

注 意

- **安定した NTP 運用を行うために、NTP コンフィグファイルには信頼できる NTP サーバを複数台（3 台以上）指定してください。**

ntpdate コマンドが参照する NTP サーバの設定

ntpd 起動時に実行される ntpdate コマンドが参照する NTP サーバの IP アドレスを、step-tickers ファイル（/etc/ntp/step-tickers）に以下のように設定します。

```
<信頼できるNTPサーバのIPアドレス-1>
<信頼できるNTPサーバのIPアドレス-2>
<信頼できるNTPサーバのIPアドレス-3>
```

注 意

- **NTP コンフィグファイルに指定した NTP サーバと同一の NTP サーバを指定してください。**
- **Step モードで運用し、かつ ntpd 起動時の ntpdate コマンドの実行を抑止したい場合は、本設定は実施しないでください。**

その他の設定項目については、「6.4 NTP 運用の設定」（→ P.130）を参照してください。

3 **drift ファイル（デフォルトでは /var/lib/ntp/drift）の内容を確認し、値が 500 または -500 となっている場合は、値を 0 に設定してください。**

設定方法については、「■ ntpd 起動前の drift 値の確認」（→ P.139）を参照してください。

4 **NTP サービスを起動します。以下のコマンドを実行してください。**

```
# /sbin/service ntpd start
```

以下の出力結果が得られた場合、NTP サービスの起動が成功したことを示します。

```
Starting ntpd:                        [  OK  ]
```

5 **上位 NTP サーバとの同期状態を確認します。以下のコマンドを実行してください。**

```
# /usr/sbin/ntpq -np
```

以下のような出力結果が得られます。IP アドレスの一番左に「*」が付いている場合、そのサーバと同期していることを表します。NTP サーバとの同期がとれていることを確認してください。

```
# /usr/sbin/ntpq -np
   remote           refid          st t when poll reach delay  offset   jitter
==============================================================================
*xxx.xxx.xxx.xxx  xxx.xxx.xxx.xxx  5 u  59    64   377  0.395 -17.244  1.659
+xxx.xxx.xxx.xxx  xxx.xxx.xxx.xxx  5 u  55    64   377  0.434  3.688   3.544
+xxx.xxx.xxx.xxx  xxx.xxx.xxx.xxx  5 u  59    64   377  0.379 -24.701  6.299
```

参 考

- **ntpd が起動して同期状態に至るまでは、およそ 5 ～ 6 分の時間がかかります。**

6 システム起動時に NTP サービスを自動起動の設定状況を確認します。以下のコマンドを実行してください。

```
# /sbin/chkconfig --list ntpd
```

以下は、出力の一例です。

```
ntpd   0:off  1:off  2:on  3:on  4:on  5:on  6:off
```

ここで、0 から 6 までの数字は、システム起動時のランレベルを意味します。
on/off は、そのランレベルでシステムが起動された場合のサービス自動起動の有無を表します。
設定を変更する場合は、目的のランレベルを指定して、以下のコマンドを実行してください。
・NTP サービスの自動起動を行う場合

```
# /sbin/chkconfig --level <ランレベル> ntpd on
```

・NTP サービスの自動起動を行わない場合

```
# /sbin/chkconfig --level <ランレベル> ntpd off
```

なお、NTP サービスの自動起動設定の状況確認において、上記の例に示す出力が得られなかった場合は、NTP サービスが自動起動の管理対象外になっている可能性があります。この場合、NTP サービスを自動起動の対象とするために、以下のコマンドを実行してください。

```
# /sbin/chkconfig --add ntpd
```

このあと、再度 NTP サービスの自動起動設定の状況確認を実施してください。

6.4 NTP 運用の設定

NTP 運用を行うための設定ファイルおよびコマンドについて説明します。

- 「6.4.1 /etc/ntp.conf」（→ P.130）
- 「6.4.2 /etc/sysconfig/ntpd」（→ P.136）
- 「6.4.3 /etc/ntp/step-tickers」（→ P.138）
- 「6.4.4 /var/lib/ntp/drift」（→ P.138）

6.4.1 /etc/ntp.conf

ntpd の設定を行うファイルです。設定を変更したあと、ntpd の再起動により設定が反映されます。/etc/ntp.conf では、参照サーバの設定、ローカルクロックの stratum の設定、アクセス制限の設定を行うことができます。

■ 参照サーバの設定

参照する上位 NTP サーバの設定を行います。以下の書式に従い、記述してください。複数の NTP サーバの設定が可能です。しかし、正確な運針をしない NTP サーバが設定された場合、正しい時刻同期の動作保証はできなくなります。ローカルクロックを設定する場合は、「127.127.1.0」を設定してください。また、prefer を設定することにより、優先する上位 NTP サーバを指定することも可能です。ただし、上位 NTP サーバとの時刻差などにより、設定した上位 NTP サーバが優先されない場合もあります。

● 書式

```
server <IPアドレス/ホスト名> [prefer]
```

● 設定例

```
server 192.168.1.2 prefer
server 0.pool.ntp.org
```

■ stratum の設定

ローカルクロックの stratum を設定し、独立した NTP サーバであることを意味します。参照する NTP サーバとしてローカルクロックを設定する場合は、以下の書式に従って記述してください。ここで設定した任意の stratum は、上位 NTP サーバの stratum がローカルクロックよりも大きくなった場合に優先して参照されます。

● 書式

```
fudge  127.127.1.0 stratum  <階層>
```

● 設定例

```
fudge  127.127.1.0 stratum  10
```

■ peer の設定

同一 Stratum 階層内の NTP サーバ間で、互いに同期しあう設定を行います。複数の peer を設定することができます。また、prefer を設定することにより、優先する peer を指定することもできます。

● 書式

```
peer  <IPアドレス|ホスト名>  [prefer]
```

注 意

- **以下の** NTP **サーバを** peer **として設定は、時刻ずれが発生する要因となります。設定しないよう注意してください。**
  - **・異なる** Stratum **階層に属する** NTP **サーバ**
  - **・同一の上位** NTP **サーバを参照している** NTP **サーバ**

● 設定例

xxx.co.jp を上位 Stratum に持つ host1 と、yyy.ne.jp を上位 Stratum に持つ host2 を peer 設定した場合の例を以下に示します。

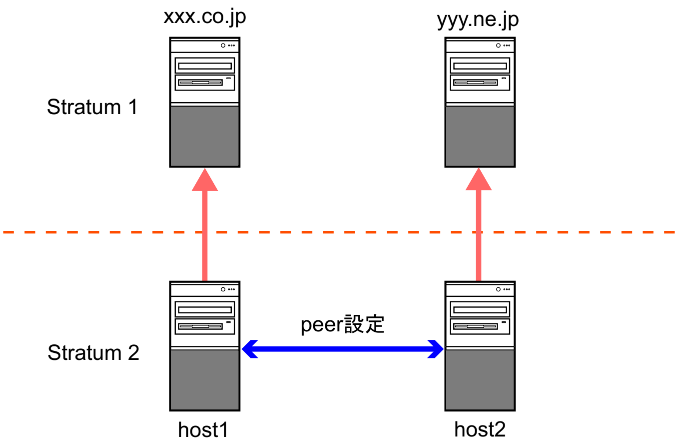


図 6 - 8：peer 設定構成例

- /etc/ntp.conf（host1）

```
server xxx.co.jp
peer   host2
```

- /etc/ntp.conf（host2）

```
server yyy.ne.jp
peer   host1
```

■ アクセス制御の設定

アクセス制御の設定を行います。アクセス制御を行う場合は、以下の書式に従って記述してください。restrict では、最後に書かれた設定が一番強くなります。そのため最初にすべてを拒否し、そのあとに許可する設定の順に記述してください。

また、アドレス指定の際の優先度は、IP アドレス→ネットワーク→ default の順になっています。default 指定は、IP アドレスや、ネットワーク指定が適用されないアドレスからの要求に対して使用されます。

● 書式

- IPv4 アドレスによる設定

```
restrict <IPアドレス/ネットワーク/default> [mask <サブネットマスク>]
[<コントロール>…]
```

- IPv6 アドレスによる設定

```
restrict -6 <IPv6アドレス/ネットワーク/default> [mask <サブネットマスク>]
[<コントロール>…]
```

表 6 - 3：[コントロール] の設定

コントロール	説明
ignore	すべての NTP に関する要求を無視します。
nomodify	時刻の問い合わせに応答しますが、サーバの設定変更や状態の確認／変更に使用されるパケットは無視します。
notrust	NTP 認証を行っていない状態での時刻参照を許可しません。
nopeer	指定ホストと相互に同期しません。
noquery	NTP の実装に依存する時刻の問い合わせを無視します。
notrap	リモートロギングに使用される種類のパケットに対してサービスを提供しません。
指定なし	NTP に関する要求のすべてを許可します。

● 設定例

```
restrict 192.168.1.0 mask 255.255.255.0 nomodify notrap
```

■ drift ファイルのパスの設定

drift ファイルの場所を設定します。以下の書式に従って絶対パスで設定してください。drift ファイルの説明は、「6.4.4 /var/lib/ntp/drift」（→ P.138）を参照してください。

● 書式

```
driftfile   <drift ファイルの絶対パス>
```

● 設定例

```
driftfile /var/lib/ntp/drift
```

■ peerstats/loopstats の設定

peerstats/loopstats ファイルを格納するディレクトリを設定します。以下の書式に従って絶対パスで設定してください。peerstats/loopstats の詳細については、「6.5.2 NTP 運用の稼動状況の確認方法」（→ P.140）を参照してください。

なお、ntpd には本ログファイルを定期的に自動で削除する機能はありません。消費するファイルサイズは 1 年で約 300MB 程度ですが、ディスク容量の圧迫の恐れがありますので、定期的に圧縮または削除してください。

注 意

- **消費するファイルサイズは、設定する NTP サーバの数に影響します。上記の例は 3 台の NTP サーバを設定した場合です。**
  ファイルを自動的に削除させる例については、「● 参考」（→ P.135）を参照してください。

● 書式

```
statsdir  ＜ログファイル格納ディレクトリのパス＞
filegen ＜ファイル生成セット名＞ [file ＜ファイル名＞] [type ＜タイプ＞]
[link|nolink] [enable|disable]
```

statsdir ＜ログファイル格納ディレクトリのパス＞の設定

統計情報のログファイルを格納するディレクトリを設定することができます。
なお、設定するディレクトリは ntp ユーザーがアクセスできる必要があります。

filegen ＜ファイル生成セット名＞の設定

採取したい統計情報のファイル生成セットを設定することができます。
＜ファイル生成セット名＞では、以下を設定することができます。

表 6 - 4：＜ファイル生成セット名＞の設定

ファイル生成セット名	説明
peerstats	同期先との時刻差などの統計情報を出力します。
loopstats	自ホストのクロック更新に関する情報を出力します。

[file ＜ファイル名＞] の設定

作成するファイル名を設定します。

[type ＜タイプ＞] の設定

＜タイプ＞の設定により、ファイルを分割して保存することができます。
＜タイプ＞では、以下を設定することができます。

表 6 - 5：＜タイプ＞の設定

タイプ	説明	ファイル名
none	ファイルを分割しません。	*＜ファイル名＞* 例）peerstats
pid	ntpd のプロセス ID ごとにファイルを作成します。	*＜ファイル名＞.＜pid ＞* 例）peerstats.9610
day	1 日ごとにファイルを作成します。[注 1]	*＜ファイル名＞.＜yyyymmdd ＞* 例）peerstats.20091027
week	1 週間ごとにファイルを作成します。[注 2]	*＜ファイル名＞.＜yyyy ＞*W*＜num ＞* 例）peerstats.2009W1
month	1ヶ月ごとにファイルを作成します。	*＜ファイル名＞.＜yyyymm ＞* 例）peerstats.200910
year	1 年ごとにファイルを作成します。	*＜ファイル名＞.＜yyyy ＞* 例）peerstats.2009
age	ntpd が 24 時間稼働するごとにファイルを作成します。	*＜ファイル名＞.a ＜sec ＞* 例）peerstats.a00000000

［注 1］：UTC 時刻の 0 時 0 分にファイルが作成されるため、日本時間の 9 時 0 分（+9 時間）に新しいファイルが作成されます。
［注 2］：年の初めから 7 日おきを一週間とします。

[link | nolink] の設定

link を設定した場合、file で指定したファイル名（サフィックスを除く）で現在のファイルのハードリンクが作成されます。これにより、現在のファイルに一定の名前でアクセスすることができます。nolink を設定した場合、ハードリンクは作成されません。
本設定を省略した場合、link の動作となります。

[enable | disable] の設定

情報採取機能を有効または無効にします。
本設定を省略した場合、enable（有効）の動作となります。

● 設定例

```
statsdir  /var/log/ntp/
filegen peerstats file peerstats type month link enable
filegen loopstats file loopstats type month link enable
```

● 参考

peerstats/loopstats の設定手順および peerstats/loopstats ファイルを定期的に削除する手順について説明します。本例は、ご利用環境に合わせて参考としてご利用ください。

1. **/etc/ntp.conf に peerstats/loopstats の設定を行います。**

実行方法

```
# vi /etc/ntp.conf
```

設定例

```
statsdir  /var/log/ntp/
filegen peerstats file peerstats type month link enable
filegen loopstats file loopstats type month link enable
```

上記の設定の場合、動作は以下のようになります。
・peerstats/loopstats ファイルを /var/log/ntp ディレクトリの配下に格納する。
・1ヶ月ごとに peerstats/loopstats ファイルを作成する。
・現在の peerstats/loopstats ファイルのハードリンクを作成する。

2. **peerstats/loopstats ファイルを格納するディレクトリを作成します。**

```
# mkdir /var/log/ntp
```

3. **手順 2 で作成したディレクトリの所有者とグループを ntp に変更します。**

```
# chown ntp.ntp /var/log/ntp
```

4. **ntpd を再起動します。**

```
# service ntpd restart
```

5　/etc/cron.monthly **ディレクトリの配下に、**peerstats/loopstats **ファイルを定期的に削除するシェルスクリプトを作成します。**

以下は、6 世代を超えたファイルを日付の古いものから順に削除するスクリプトの例です。

```
#!/bin/bash
STATSDIR="/var/log/ntp"      # statsdirで設定したディレクトリ
PFILE="peerstats"            # peerstatsのファイル名
LFILE="loopstats"            # loopstatsのファイル名
AGE=6                        # 保存する世代数

while [ `/bin/ls $STATSDIR/$PFILE.* | wc -l` -gt $AGE ];
do
    DELLIST=`/bin/ls $STATSDIR/$PFILE.* | /bin/sort | /usr/bin/head -n 1`
    /bin/rm -f $DELLIST
    /bin/sleep 1
done
while [ `/bin/ls $STATSDIR/$LFILE.* | wc -l` -gt $AGE ];
do
    DELLIST=`/bin/ls $STATSDIR/$LFILE.* | /bin/sort | /usr/bin/head -n 1`
    /bin/rm -f $DELLIST
    /bin/sleep 1
done
```

6　**手順 5** **で作成したシェルスクリプトに実行権限を付与します。**

```
# chmod +x <ファイル名>
```

7　crond **を再起動します。**

```
# service crond restart
```

6.4.2　/etc/sysconfig/ntpd

ntpd 起動時に引数を ntpd へ渡すために使用されるファイルです。設定を変更したあと、ntpd の再起動により設定が反映されます。/etc/sysconfig/ntpd では、ntpd の起動オプションの設定ができます。ntpd は参照する NTP サーバと同期をとるためにシステムクロックを設定して管理します。

■ OPTIONS

ntpd の動作オプションを設定することができます。

● 書式

```
OPTIONS="[ntpdの動作オプション]"
```

● 設定例 1：ゆっくりとした時刻補正をする場合（Slew モード）

[-x] オプション（下線箇所）を追加してください。Step モード／ Slew モードについては、「6.2.4 時刻補正の方法」（→ P.118）を参照してください。

```
OPTIONS="-x -u ntp:ntp -p /var/run/ntpd.pid"
```

● 設定例 2：デバッグモードの場合

「-d -D 10 -l ログ出力ファイル名」（下線箇所）を追加してください。障害発生時などにデバッグログを出力するために設定します。通常運用時は、設定の必要はありません。

```
OPTIONS="-u ntp:ntp -p /var/run/ntpd.pid -d -D 10 -l /tmp/ntpdebug.log"
```

■ SYNC_HWCLOCK

ntpd 起動時にハードウェアクロックとの時刻同期を行うかを選択することができます。
デフォルトは「no」になっています。

● 書式

```
SYNC_HWCLOCK=<yes / no>
```

● 設定例

```
SYNC_HWCLOCK=yes
```

■ NTPDATE_OPTIONS

ntpd 起動時に実行される ntpdate コマンドの動作を設定するためのコマンドラインオプションを記述することができます。主なオプションを以下に示します。

表 6 - 6：ntpdate コマンドのオプション

オプション	動作説明
-B	リファレンスサーバとの時間差にかかわらず、常に Slew モードで時刻補正を行います。
-b	リファレンスサーバとの時間差にかかわらず、常に Step モードで時刻補正を行います（省略可）。

● 書式

```
NTPDATE_OPTIONS="[ntpdateコマンドの動作オプション]"
```

● 設定例

Slew 調整で実行する場合

```
NTPDATE_OPTIONS="-B"
```

6.4.3 /etc/ntp/step-tickers

ntpd 起動前の ntpdate コマンドが時刻補正を行う際に参照するファイルです。ntpd は、このファイルに記述されている NTP サーバに問い合わせて時刻補正を行います。
/etc/ntp/step-tickers がない、またはファイルに NTP サーバの記述がない場合、以下のように ntpd の起動時の動作が Step モードと Slew モードで異なります。

- Step モードでの起動時
  ntpdate による時刻補正を行いません。
- Slew モードでの起動時
  /etc/ntp.conf の server ディレクティブに記載されている NTP サーバに問い合わせて時刻補正を行います。

6.4.4 /var/lib/ntp/drift

NTP クライアントの ntpd が NTP サーバと時刻同期する際に生じた時刻差などから算出された補正値を記録するファイルです。NTP クライアントの ntpd は、起動時に /var/lib/ntp/drift にある周波数に基づいて時刻補正を開始します。
最適な drift 値を使用した場合、NTP クライアントは ntpd を起動した直後から、NTP サーバとほぼ同じスピードで時刻補正を開始します。

6.5 NTP 運用時の対応

ここでは、NTP 運用時の考慮点と、稼動状況の確認方法について説明します。

- 「6.5.1 NTP 運用時の考慮点」（→ P.139）
- 「6.5.2 NTP 運用の稼動状況の確認方法」（→ P.140）

6.5.1 NTP 運用時の考慮点

■ NTP サーバの階層を考慮した NTP サーバの起動順序

NTP クライアントは、Stratum1、またはさらに上位の NTP サーバと同期中である NTP サーバとのみ同期できます。上位 NTP サーバと NTP クライアントを同時に起動するなどの運用を行っている場合、NTP クライアントが起動時に時刻同期することができず、時刻補正に問題が発生する可能性があります。そのため、NTP 運用を行う場合には、上位 NTP サーバから順番に起動するようにサーバの起動タイミングに注意してください。
ntpd が起動して同期状態になるまで、およそ 5 ～ 6 分かかります。NTP クライアントを起動するのは、参照先 NTP サーバを起動してから 15 ～ 20 分（OS 起動時間も含めた時間）経過したあとに実施してください。

■ ntpd 起動前の drift 値の確認

ntpd を起動する前に、drift 値を確認してください。drift 値は /var/lib/ntp/drift に記載されています。drift 値は、安定した NTP 運用を行っている場合に設定されている値を維持すべきです。drift 値に、その値から極端に離れた値（drift 値の限界値は± 500）が設定されている場合には、上位 NTP サーバとの同期を妨げる原因になる可能性があります。その場合には、ntpd を起動する前に値を 0 に変更してください。実行例は以下のとおりです。

● 実行例

1 **NTP サービスを停止します。**

```
# service ntpd stop
```

2 **/var/lib/ntp/drift ファイルの値を 0 に変更します。**

```
# echo -n "0" > /var/lib/ntp/drift
```

3 **NTP サービスを起動します。**

```
# service ntpd start
```

■ ntpdate コマンドの実施時の考慮

ntpdate コマンドは、随時起動のプログラムとして時刻補正を行います。
ntpd が稼働している場合は、ntpdate コマンドを使用して時刻を補正することができません。
そのため、ntpdate を使用する場合には、ntpd をいったん停止する必要があります。

6.5.2 NTP 運用の稼動状況の確認方法

■ 上位 NTP サーバとの時刻同期状況の確認

NTP 運用を開始した際は、必ず設計したとおりに時刻同期が行われていることを確認してください。上位 NTP サーバとの時刻同期状況を確認するには、ntpq コマンドを使用します。
以下のコマンドを実行し、出力結果を確認します。

```
# /usr/sbin/ntpq -np
```

注 意

- **ntpd が起動して同期状態になるまで、およそ 5 ～ 6 分かかります。**

以下のような出力結果が得られます。出力される各行は、同期対象とする NTP サーバとの同期状況を示します。
IP アドレスの一番左に「*」が付いている場合、そのサーバと同期していることを表します。

```
 # /usr/sbin/ntpq -np
      remote           refid          st t when poll reach   delay   offset   jitter
 ==============================================================================
 *xxx.xxx.xxx.xxx  xxx.xxx.xxx.xxx  5 u   59   64  377    0.395  -17.244  1.659
 +xxx.xxx.xxx.xxx  xxx.xxx.xxx.xxx  5 u   55   64  377    0.434    3.688  3.544
 +xxx.xxx.xxx.xxx  xxx.xxx.xxx.xxx  5 u   59   64  377    0.379  -24.701  6.299
(1)       (2)              (3)         (4)(5)  (6)  (7)  (8)      (9)      (10)     (11)
```

(1) 参照している ntp サーバとの状態
　表示される記号の意味は以下のとおりです。
　"*" ：現在参照同期中のサーバ
　"+" ：クロック誤り検査に合格したサーバ
　"#" ：参照同期中だが，距離が遠いサーバ
　" " ：現在参照していないサーバ
　"x" ：クロック誤り検査で不合格になったサーバ
　"." ：参照リストから外されたサーバ
　"-" ：クラスタリング検査で参照リストから外されたサーバ

(2) ntp サーバのサーバ名（または IP アドレス）

(3) ntp サーバが参照している、さらに上位の ntp サーバ（不明の場合は 0.0.0.0）
　".INIT." や ".STEP." などと表示される場合があります。
　".INIT." や ".STEP." は、初期化中や起動処理中に表示されます。

(4) stratum 番号、参照している ntp サーバが第何階層かを表します。

(5) 階層タイプ（l:local、u:unicast、m:multicast、b:broadcast）

(6) 最後のパケットを受け取ってからの時間（単位：秒）

(7) ポーリング間隔（単位：秒）

(8) 最近 8 回のポーリング状況（8 進数表現）
ポーリング状況が、成功：1、失敗：0 のビットとした 8 進数で表示されます。
最新のポーリング状態が右端（001）になり、左にシフトするごとに過去のポーリング状態を示します。

(9) ポーリング間隔の遅延見積り（単位：ミリ秒）

(10) 階層のオフセット（単位：ミリ秒）

(11) 階層の分散（単位：ミリ秒）

■ 参照 NTP サーバとの時刻差の情報確認（peerstats）

ntp.conf の statsdir で指定したディレクトリの配下にファイルを作成することによって、同期先との時刻差などの統計情報を出力することができます。このファイルの時刻差の変化を分析することで、NTP サーバとどの程度の誤差で動作していたのかを確認することができます。ntpd の時刻補正に問題が発生した場合の調査に有効なため、本情報を採取することを推奨します。設定方法については、「6.4.1 /etc/ntp.conf」（→ P.130）を参照してください。

ファイル中のレコードの出力例と、それぞれのフィールドの説明は以下のとおりです。

```
55112 27034.251 192.168.1.10 9614 0.000870735 0.002303132 0.018681483 0.000302981
 (1)     (2)        (3)       (4)     (5)         (6)         (7)         (8)
```

(1) 日付。修正ユリウス日を示します。

(2) 時刻（単位：秒)。0 時 0 分 0 秒（UTC）からの経過秒数を示します。日本時間にするには +9 時間（+32400 秒）する必要があります。

(3) 相手 IP アドレス（127.127.1.0 はローカルクロック）

(4) 同期先の状態（16 進表示）

(5) NTP サーバとの時刻差（単位：秒）。peerstats 内の時刻差は、ntpd 内の内部時刻と上位 NTP サーバの ntpd 内部時刻との差であり、OS のシステム時計と必ずしも一致していません。

(6) 遅延時間（単位：秒）

(7) 分散値（単位：秒）

(8) RMS ジッタ（単位：秒）

上記出力例の場合、192.168.1.10 のサーバと 0.870735 ミリ秒の時刻差があることを示しています。このファイルの時刻差の変化を分析することで、NTP サーバとどの程度の誤差で動作していたのかを検証することができます。
図 6 - 9 に、NTP サーバとの時刻差の変化の例を示します。

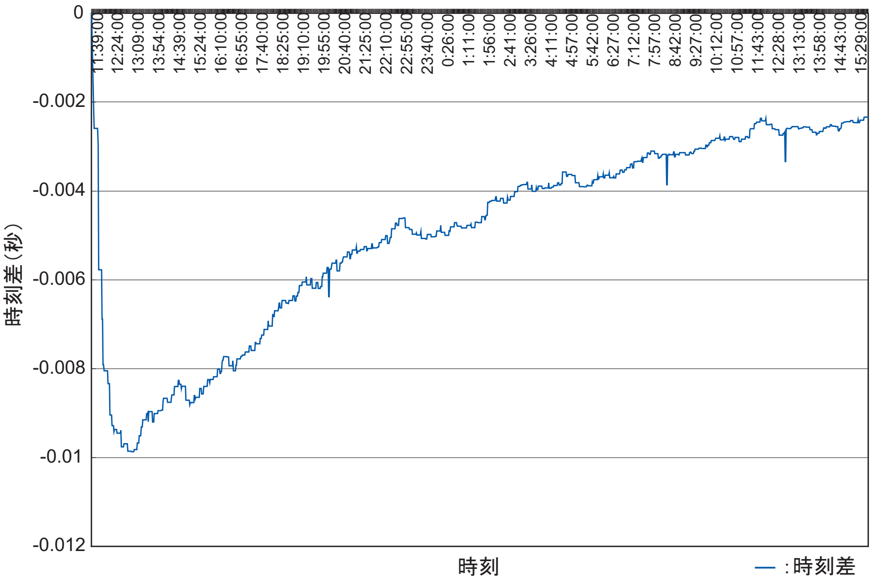


図 6 - 9：NTP サーバとの時刻差の変化の例

■ 自ホストのクロック更新に関する情報確認（loopstats）

ntp.conf の statsdir で指定したディレクトリの配下にファイルを作成することによって、自ホストのクロック更新に関する情報を出力することができます。なお、loopstats の情報は NTP サーバ（またはローカルクロック）と同期状態にある場合のみ出力されます。
ntpd の時刻補正に問題が発生した場合の調査に有効なため、本情報を採取することを推奨します。設定方法については、「6.4.1 /etc/ntp.conf」（→ P.130）を参照してください。

ファイル中のレコードの出力例と、それぞれのフィールドの説明は以下のとおりです。

```
55139 21009.957 -0.000264663 -65.864 0.000095931 0.001398 6
 (1)     (2)         (3)        (4)        (5)        (6)   (7)
```

(1) 日付。修正ユリウス日を示します。
(2) 時刻（単位：秒）。0 時 0 分 0 秒（UTC）からの経過秒数を示します。日本時間にするには +9 時間（+32400 秒）する必要があります。
(3) NTP サーバとの時刻差（単位：秒）。loopstats 内の時刻差は、ntpd 内の内部時刻と上位 NTP サーバの ntpd 内部時刻との差であり、OS のシステム時計と必ずしも一致していません。
(4) drift 値（単位：PPM（parts-per-million））
(5) RMS ジッタ（単位：秒）
(6) クロック安定度（単位：PPM（parts-per-million））
(7) 自ホストのクロック更新定数（polling interval）。2 のベキ指数が表示されます。

6.6 うるう秒への対応

うるう秒への対応方法について説明します。

6.6.1 うるう秒とは

現在、時刻は地球の自転に基づく天文時と、原子の振動を利用した原子時計によって決められています。これらの原子時計と天文時に基づく時刻との間では、少しずつずれが生じます。そのため、原子時計に基づく時刻を天文時と 0.9 秒以上ずれないように、1 秒を挿入または削除して時刻の調整を行います。この操作をうるう秒といいます。これまでに実施されたうるう秒は、すべて 1 秒を挿入する操作です。

NTP による時刻補正を行っている場合、うるう秒の情報は NTP パケットに含まれる Leap Indicator によって下位の NTP クライアントに伝達されます。

6.6.2 うるう秒対応

うるう秒挿入時のシステム時刻の時刻遷移は、zoneinfo の使用／ NTP の運用の有無／時刻補正モード（Step モード／ Slew モード）によって異なります。ここでは条件別に、システム時刻の時刻遷移の違いと、それにより与えられるシステム影響と対処について説明します。

- うるう秒対応の zoneinfo を使用している場合（ntp の使用有無の条件なし）
  「■ うるう秒対応の zoneinfo を使用している場合」（→ P.144）を参照してください。
- うるう秒未対応の zoneinfo を使用している場合
  - ntp 運用なし
    「● NTP 運用していない場合」（→ P.145）を参照してください。
  - ntp 運用あり
    以下を参照してください。
    「● NTP による時刻補正（Step モード）をしている場合」（→ P.145）
    「● NTP による時刻補正（Slew モード）をしている場合（ntp-4.2.2p1-8.el5 以前の場合）」（→ P.146）
    「● NTP による時刻補正（Slew モード）をしている場合（ntp-4.2.2p1-9.el5 以降の場合）」（→ P.146）

参考

うるう秒対応／未対応の判断方法

- md5sum コマンドにより適用している zoneinfo のチェックサム値を調べることで確認できます。/etc/localtime のチェックサム値が以下のいずれの zoneinfo と同じかを確認してください。

表 6 - 7：チェックサム値

zoneinfo	ファイル名
うるう秒未対応	/usr/share/zoneinfo/Asia/Tokyo
うるう秒対応	/usr/share/zoneinfo/right/Asia/Tokyo

コマンドの実行例は以下のとおりです。

```
# md5sum /etc/localtime
8470e9e89799169fde673f92103a72f8 /etc/localtime
チェックサム値

# md5sum /usr/share/zoneinfo/Asia/Tokyo
8470e9e89799169fde673f92103a72f8 /usr/share/zoneinfo/Asia/Tokyo
チェックサム値（うるう秒未対応zoneinfo）

# md5sum /usr/share/zoneinfo/right/Asia/Tokyo
3489e452e74de59fse645f9fj76a7sd4 /usr/share/zoneinfo/right/Asia/Tokyo
チェックサム値（うるう秒対応zoneinfo）
```

この場合、/etc/localtime がうるう秒未対応の zoneinfo とチェックサムが同じため、うるう秒未対応の zoneinfo を使用したシステムとなります。

■ うるう秒対応の zoneinfo を使用している場合

時刻遷移

NTP 運用の有無にかかわらず、うるう秒対応の zoneinfo を使用した場合、うるう秒挿入時には以下のように 8 時 59 分 60 秒を刻みます。

```
日本標準時        システム時刻
08:59:59         08:59:59
08:59:60         08:59:60
09:00:00         09:00:00
```

影響

うるう秒挿入時に 8 時 59 分 60 秒を刻むため、60 秒を認識できないミドルウェアやアプリケーションにおいて動作や表示の不具合などの影響が発生する可能性があります。

対処

うるう秒対応の zoneinfo は使用しないでください。/etc/localtime がうるう秒未対応の zoneinfo（/usr/share/zoneinfo/Asia/Tokyo）とチェックサム値が異なる場合、/etc/localtime をうるう秒未対応の zoneinfo で上書きしてください。

実行方法

```
# cp /usr/share/zoneinfo/Asia/Tokyo /etc/localtime
```

■ うるう秒未対応の zoneinfo を使用している場合

● NTP 運用していない場合

時刻遷移

うるう秒として特別な処理はないため、うるう秒挿入後は以下のように 1 秒ずれた時刻遷移になります。

```
日本標準時        システム時刻
08:59:59          08:59:59
08:59:60          09:00:00
09:00:00          09:00:01
```

影響

日本標準時と 1 秒ずれます。

● NTP による時刻補正（Step モード）をしている場合

時刻遷移

うるう秒挿入時に、再度 8 時 59 分 59 秒を刻みます。

```
日本標準時          システム時刻
08:59:59.999        08:59:59.999
08:59:60.000        09:00:00.000
08:59:60.001        08:59:59.001  ←1秒の逆進が発生
08:59:60.002        08:59:59.002
     ～                  ～
08:59:60.999        08:59:59.999
09:00:00.000        09:00:00.000
```

影響

うるう秒挿入時に再度 8 時 59 分 59 秒を刻むため、1 秒の時刻の逆進が発生します。時刻が逆進することにより、ミドルウェアやアプリケーションにおいて不具合が発生する可能性があります。

対処

時刻の逆進を許容しないシステムの場合は、Slew モードを使用してください。設定方法については、「6.4.2 /etc/sysconfig/ntpd」（→ P.136）を参照してください。なお、Slew モードでの対処については、「● NTP による時刻補正（Slew モード）をしている場合（ntp-4.2.2p1-8.el5 以前の場合）」（→ P.146）または「● NTP による時刻補正（Slew モード）をしている場合（ntp-4.2.2p1-9.el5 以降の場合）」（→ P.146）を参照してください。

● NTP による時刻補正（Slew モード）をしている場合（ntp-4.2.2p1-8.el5 以前の場合）

時刻遷移

うるう秒挿入時に、再度 8 時 59 分 59 秒を刻みます。

日本標準時	システム時刻	
08:59:59.999	08:59:59.999	
08:59:60.000	09:00:00.000	
08:59:60.001	08:59:59.001	←1秒の逆進が発生
08:59:60.002	08:59:59.002	
～	～	
08:59:60.999	08:59:59.999	
09:00:00.000	09:00:00.000	

影響

うるう秒挿入時に再度 8 時 59 分 59 秒を刻むため、1 秒の時刻の逆進が発生します。時刻が逆進することにより、ミドルウェアやアプリケーションにおいて不具合が発生する可能性があります。

対処

以下のどちらかの対処を行ってください。

a) うるう秒挿入までにシステム（OS）を停止し、うるう秒挿入後にシステムを起動する。

b) うるう秒挿入の前日 9 時 0 分以前に ntpd を停止し、うるう秒挿入後に ntpd を起動する。

なお、対処 b) の場合、ntpd 停止中はシステムクロックに基づいて時刻を刻むため、時刻精度が低下します。クラスタなどを使用している場合にはシステムに影響が出る可能性があるため、ntpd 停止中は ntpdate コマンドによる時刻調整を行う必要があります。対処の詳細については、「6.6.3 対処詳細」（→ P.147）を参照してください。

● NTP による時刻補正（Slew モード）をしている場合（ntp-4.2.2p1-9.el5 以降の場合）

時刻遷移

うるう秒挿入時も通常どおりの時刻を刻むため、うるう秒挿入時点では 1 秒のずれが発生します。その後、NTP によって 1 秒のずれは調整されます。

日本標準時	システム時刻
08:59:59	08:59:59
08:59:60	09:00:00
09:00:00	09:00:01

影響

日本標準時と 1 秒ずれます。

対処

NTP によって 1 秒のずれは調整されるため、対処の必要はありません。

6.6.3 対処詳細

以下の手順に従って、うるう秒が挿入される前日 9 時 0 分以前に ntpd を停止し、うるう秒の挿入以降に ntpd を起動してください。うるう秒挿入の前日 9 時 0 分以降は、ntpd は上位 NTP サーバから受け取ったうるう秒情報（Leap Indicator）をカーネルに伝達します。うるう秒情報（=01）がカーネルに伝達されると、1 秒の時刻の逆進が発生するため、ntpd を停止する必要があります。ntpd 運用を停止期間中は、うるう秒情報をカーネルに伝達しない ntpdate コマンドによる時刻補正を行ってください。

注 意

- **以下の作業は、すべて root 権限で行ってください。**

1 ntpd の起動スクリプトを変更します。

1. **起動オプションファイルをバックアップします。**

```
# cp /etc/sysconfig/ntpd /etc/sysconfig/ntpd.org
```

2. **ntpdate のオプションを変更します。**

```
# vi /etc/sysconfig/ntpd
```

以下のように変更してください。

・変更前

```
NTPDATE_OPTIONS=""
```

・変更後

```
NTPDATE_OPTIONS="-B"
```

2 ntpd を停止します。

うるう秒挿入の前日 9 時 0 分までに ntpd を停止します。

1. **ntpd を停止します。**

```
# service ntpd stop
```

以下のように表示されることを確認してください。

```
Shutting down ntpd:      [ OK ]
```

2. **ntpd が停止していることを確認します。**

```
# service ntpd status
```

以下のように表示されることを確認してください。

```
ntpd is stopped
```

3 ntpdate コマンドによる時刻補正を開始します。

以下のコマンドを定期的に実行してください。

```
# ntpdate -B  <NTPサーバのIPアドレス>
```

参 考

- 2009 年 1 月 1 日の対応実績より、実行間隔は 600 秒を推奨します。

4 ntpd を起動します。

うるう秒挿入以降に、上位 NTP サーバとの offset の値（下線箇所）が 0.1sec 以下であることを確認し、ntpd を起動してください。ntpd を起動したあとに手順 3 は行わないでください。

1. **以下のコマンドを実行します。**

```
# ntpdate -q  <NTPサーバのIPアドレス>
```

offset の値（下線箇所）が 0.1sec 以下であることを確認します。

```
server xxx.xxx.xxx.xxx, stratum 6, offset 0.065965, delay 0.02583
ntpdate[<PID>]: adjust time server xxx.xxx.xxx.xxx offset 0.065965 sec
```

2. **ntpd を起動します。**

```
# service ntpd start
```

以下のように表示されることを確認してください。

```
ntpd: Synchronizing with time server:        [ OK ]
Starting ntpd:                               [ OK ]
```

3. **ntpd が起動していることを確認します。**

```
# service ntpd status
```

以下のように表示されることを確認してください。

```
ntpd (pid <PID>) is running...
```

5 ntpd スクリプトを復元します。

手順 1 でバックアップした起動スクリプトに戻します。

```
# mv /etc/sysconfig/ntpd.org /etc/sysconfig/ntpd
```

7

第 7 章

システム構築：OS 環境構築

この章では、OS 環境構築の方法について説明します。

7.1 カスタマーポータルへのサブスクリプション登録

RHEL をご使用いただくためには、カスタマーポータルへのサブスクリプションの登録が必要です。
サブスクリプションの登録をしていただくことにより、インストールイメージ（ISO ファイル形式）やバグ／セキュリティ修正の入手など、Red Hat 社のサポートサービスを受けることができます。

• 「7.1.1 サブスクリプションの登録／更新」（→ P.150）

7.1.1 サブスクリプションの登録／更新

サブスクリプションの登録／更新方法について説明します。

注 意

- **サブスクリプションの登録の前に、Red Hat 社の「Red Hat Agreements」サイト（http://www.redhat.com/licenses/）を確認してください。**
  「Japan」をクリックすると、日本語の PDF 契約書を参照できます。

■ サブスクリプションの登録

初めてカスタマーポータルを利用する場合は、まずカスタマーポータルのアカウント作成が必要です。カスタマーポータルのアカウント作成手順については、Red Hat 社の以下のサイトを参照してください。

http://www.jp.redhat.com/FAQ/regist.html#1

作成したアカウントでカスタマーポータルにログインして、サブスクリプションの登録を行います。サブスクリプション登録の手順については、Red Hat 社の以下のサイトを参照してください。

http://www.jp.redhat.com/FAQ/regist.html#2

■ サブスクリプションの更新

サブスクリプションには有効期限（1 年または複数年）があり、利用期限を過ぎる前に更新手続きをする必要があります。サブスクリプションの更新手順については、Red Hat 社の以下のサイトを参照してください。

http://www.jp.redhat.com/FAQ/regist.html#4

7.2 RHEL のインストール

RHEL のインストール方法について説明します。

- 「7.2.1 RHEL のインストール方法」(→ P.151)
- 「7.2.2 RHEL のインストール時の注意事項」(→ P.151)

7.2.1 RHEL のインストール方法

RHEL のインストールは、構築する環境により異なり、PRIMEQUEST 1000 シリーズに添付されている SVIM を使用してインストールする方法と、SVIM を使用しないでインストールする方法があります。構築する環境ごとのインストール方法、使用するツール、および参照ドキュメントは、表 7 - 1 のとおりです。

表 7 - 1：インストール方法と参照ドキュメント

構築する環境		インストール方法	参照ドキュメント
ネイティブ OS		SVIM を使用する	『導入マニュアル』
Xen	管理 OS	SVIM を使用する	『導入マニュアル』 『RHEL5-Xen 仮想マシン機能ユーザーズマニュアル』
	ゲスト OS	SVIM を使用しない	『RHEL5-Xen 仮想マシン機能ユーザーズマニュアル』
KVM	ゲスト OS	SVIM を使用しない	『Linux ユーザーズマニュアル』(本書)
VMware	ゲスト OS	SVIM を使用しない	『Linux ユーザーズマニュアル』(本書)

SVIM を使用して RHEL をインストールする場合は、インストール作業を開始する前に参照ドキュメントを確認し、SVIM で必要となる入力項目をあらかじめ決定してください。

7.2.2 RHEL のインストール時の注意事項

RHEL のインストール時の注意事項について説明します。

■ SAN Boot 環境の設定

SAN Boot 環境を構築するためには、RHEL のインストール前にハードウェアの環境設定が必要です。設定方法については、『導入マニュアル』を参照してください。

参 考

- **SAN Boot 環境とは、サーバ筐体に内蔵されたディスクではなく、ファイバーチャネルで接続した外部の SAN ストレージ装置のディスクからシステムを起動する環境です。**

■ ファイバーチャネルカード用ドライバ使用時の設定

ファイバーチャネルカードを使用する場合は、インストール後に、構築する SAN 環境に基づいたドライバオプションの設定が必要となります。そのため、インストール前に SAN 環境を確認し、以下のドライバオプションの内容をあらかじめ決定しておいてください。
決定したドライバオプションは、インストール後に /etc/modprobe.conf ファイルの末尾に追記します。なお、ドライバオプションは 1 行で記述してください。

● 転送速度 lpfc_link_speed

ファイバーチャネル経由でのアクセスにおける転送速度を指定します。指定値はシステムで一意に決めておく必要があり、以下のように算出します。

1. **搭載されたすべてのファイバーチャネルカードの最大転送速度を調べて、そのうちの最大値を求めます（=A とする）。**

2. **接続されたすべてのディスクアレイ装置の最大転送速度を調べて、そのうちの最小値を求めます（=B とする）。**

3. **A、B のうち、小さい方の値を lpfc_link_speed とします。**

● 接続形態 lpfc_topology

SAN 環境の接続形態を指定します。

- Fabric 接続（ファイバーチャネルスイッチと接続）する場合

```
lpfc_topology=2
```

- Arbitrated Loop 接続（ループ接続）する場合

```
lpfc_topology=4
```

● 記述例

以下の接続形態を持つシステムの場合のドライバオプション記述例です。

- 8Gb/s のファイバーチャネルカードと 4Gb/s のディスクアレイ装置を接続
- 4Gb/s のファイバーチャネルカードと 2Gb/s のディスクアレイ装置を接続
- Fabric 接続

この場合、A=8、B=2 となり、lpfc_link_speed は 2 となります。Fabric 接続なので lpfc_topology は 2 となります。したがって、ドライバオプションは以下のようになります。

```
options lpfc lpfc_link_speed=2 lpfc_topology=2
```

注 意

- **本設定を行わない場合、正常に接続先デバイスが認識できないことがあります。**
- **ドライバ設定の変更後は、必ず初期 RAM ディスクを再作成してください。**

7.3 環境設定

RHEL のインストール後、以下の設定を行います。

- **「7.3.1 デバイス名の変更防止の設定」（→ P.153）**
- **「7.3.2 時刻補正の導入（NTP）」（→ P.153）**

7.3.1 デバイス名の変更防止の設定

ハードウェア故障などによるデバイス名ずれの防止設定を実施します。
設定方法については、「4.2 設計・導入時の考慮」（→ P.62）を参照してください。

7.3.2 時刻補正の導入（NTP）

NTP（Network Time Protocol）による時刻補正の設定を実施します。設定方法については、「第 6 章 システム設計：時刻補正の導入（NTP）」（→ P.110）を参照してください。

7.4 レスキューモードの起動（トラブル対応時）

レスキューモードの概要と使い方について説明します。

- 「7.4.1 レスキューモードの概要」（→ P.154）
- 「7.4.2 レスキューモードの準備」（→ P.154）
- 「7.4.3 レスキューモードの起動方法」（→ P.155）
- 「7.4.4 レスキューモードの終了方法」（→ P.156）

7.4.1 レスキューモードの概要

OS の環境構築の過程で、ハードディスクからシステムの起動ができなくなった場合や、ハードディスクを使わずにシステムを起動する必要がある場合、RHEL のディストリビューションメディアを用いて、最小限の RHEL 環境を起動することができます。これをレスキューモードと呼びます。
レスキューモードは、主に以下の場合に利用します。

- ハードディスクにインストールしたシステムが起動できない場合
- システムボリュームをオフラインでバックアップ／リストアする場合

レスキューモードの起動は、RHEL のディストリビューションメディアの 1 枚目（ブータブルメディア）を用いて行います。
レスキューモードでは最小限のシステムで起動されるため、接続されていても認識されないハードウェアデバイスが存在する場合があります。認識されないハードウェアデバイスをシステムに認識させるためには、そのハードウェアデバイスに対するデバイスドライバの組み込みを行う必要があります。

7.4.2 レスキューモードの準備

レスキューモードで起動する前の準備について説明します。

● RHEL ディストリビューションのブータブルメディアの準備

RHEL ディストリビューションのブータブルメディアの 1 枚目を用意します。

● 起動環境の設定

準備したブータブルメディアから起動ができるようにハードウェアを設定します。

注 意

- **ブータブルメディアから起動ができるようにハードウェアを設定する方法については、お使いのハードウェアのマニュアルを参照してください。**

7.4.3 レスキューモードの起動方法

レスキューモードでの起動方法について説明します。

1 **ブータブルメディアをセットして、サーバを起動します。**

2 **boot プロンプトで、「linux rescue」と入力します。**

```
boot: linux rescue
```

3 **「Choose a Language」画面で［English］を選択します。**

ここで［Japanese］を選択すると、LANG=ja_JP.UTF-8 の環境となります。
一部のコマンドで画面表示が正常に行われないため、［English］を選択してください。

4 **「Keyboard Type」画面で［jp106］を選択します。**

フラットディスプレイ（PG-R1DP2）などの付属キーボードを使用する場合には、［us］を選択します。

5 **「Rescue Method」画面で［local CDROM］を選択します。**

注 意

- **サーバによっては、本画面が表示されない場合があります。表示されない場合は、****手順 8**** 以降を実施してください。**

6 **「No driver found」画面で［Select driver］を選択します。**

注 意

- **サーバによっては、本画面が表示されない場合があります。表示されない場合は、****手順 8**** 以降を実施してください。**

7 **「Select Device Driver to Load」画面で［USB Mass Storage driver for Linux (usb-storage)］を選択します。**

8 「Setup Networking」画面でネットワークの設定を行うか選択します。

ネットワークを利用しない場合は［No］を選択します。

注 意

- 「8.3.6 ファイルシステムのリモートリストア（テープ装置を使用の場合）」（→ P.186）または「8.3.7 ファイルシステムのリモートリストア（データカートリッジを使用の場合）」（→ P.189）を実施する場合は、［Yes］を選択し、ネットワークの設定を行ってください。
  設定方法は、以下のとおりです（xxxx はネットワークインターフェース名です）。
  1. 「Setup Networking」画面で［Yes］を選択します。
  2. 「Network Configuration」画面でバックアップサーバに接続されたネットワークインターフェース名を 1 つ選択したあと、［Edit］を選択します。
  3. 「Network Configuration for xxxx」画面で［Enable Ipv4 support］を選択し、［OK］を選択します。
  4. 「IPv4 Configuration for xxxx」画面で、［Manual address configuration］を選択したあと、「IP Address」と「Prefix(Netmask)」に設定する値を入力し、［OK］を選択します。または［Dynamic IP Configration(DHCP)］を選択したあと、［OK］を選択します。
  5. 「Network Configuration」画面で［OK］を選択します。
  6. 「Miscellaneous Network Settings」画面で、「Gateway」、「Primary DNS」、「Secondary DNS」に設定する値を入力し、［OK］を選択します。
     設定が不要な場合は、何も入力せずに［OK］を選択します。

9 「Rescue」画面で既存のパーティションを自動的にマウントするかどうかのメッセージが表示されますので、［Skip］を選択します。

コマンドプロンプトが表示されます。

10 ハードディスクにアクセスする場合、以下のコマンドを実行し、ハードディスクをマウントします。

例）/dev/sda2 をマウントする場合

```
# mkdir /mnt/mountpoint
# mount /dev/sda2 /mnt/mountpoint
```

7.4.4 レスキューモードの終了方法

レスキューモードを終了する場合は、コマンドプロンプトで reboot コマンドを実行してください。レスキューモードで起動したシステムが終了し、システムが再起動されます。

```
# reboot
```

8

第 8 章

運用／保守：バックアップ／リストア

この章では、資産を確実に保護するために必要なバックアップ／リストアについて説明します。

8.1 バックアップ／リストアの概要

バックアップ／リストアの概要について説明します。

- 「8.1.1 バックアップの実施方法」（→ P.158）
- 「8.1.2 バックアップの構成」（→ P.160）
- 「8.1.3 準備と環境確認」（→ P.161）

■ バックアップ／リストアの要件

お客様資産を確実に保護するためには、バックアップとリストアを対にして、バックアップ／リストア計画を立案する必要があります。
一般的にバックアップ／リストアには、以下の要件があります。

- トラブルから確実に復旧できること（確実性／安全性）
- 業務への影響を最小限に抑止すること（スピード）
- ソフト購入費、人件費を最小限に抑止すること（コスト）

注 意

- **異なる機種間でのバックアップおよびリストアは絶対に実施しないでください。また、同一機種の場合でもハードウェア構成が異なる場合には、バックアップおよびリストアは絶対に実施しないでください。システムが正常に起動できない場合があります。**
- **dump コマンドの実行時と異なる OS 版数およびアップデートシステムのレスキュー上でリストアを実施した場合、restore コマンドが失敗する可能性があります。restore コマンドを実行する際は、dump コマンド実行時と同じバージョンのインストールディスクのレスキューモードで起動してリストアを実行してください。**

8.1.1 バックアップの実施方法

バックアップの実施には、以下の 2 種類があります。業務形態に応じて計画してください。

● オンラインバックアップ

業務を続行したまま、バックアップを実行します。
オンラインバックアップを行うには、オンラインバックアップ対応製品である必要があり、運用手順が複雑になります。また、事前に業務への影響がないかどうかを検証しておく必要があります。

注 意

- **本書では、オンラインバックアップについては解説しません。**

● オフラインバックアップ

業務を停止させてから、バックアップを実行します。
運用手順は簡単ですが、業務（サービス）の停止が必要です。
業務停止時間はバックアップ容量に比例し、バックアップ対象の領域が大きいほど業務停止時間は長くなります。
なお、オフラインバックアップには、以下の 3 つのバックアップ方法があります。

- フルバックアップ
  毎回すべての領域をバックアップします。リカバリー時には、リカバリー対象日のバックアップデータを使用します。
- 差分バックアップ
  全体バックアップ時点からの変更データをバックアップとして採取します。リカバリー時には、最新の全体バックアップデータと、リカバリー対象日の差分バックアップデータを使用します。
- 増分バックアップ
  前回増分バックアップ時点からの変更データをバックアップとして採取します。リカバリー時には、最新の全体バックアップデータと、リカバリー対象日までの増分バックアップデータのすべてを使用します。

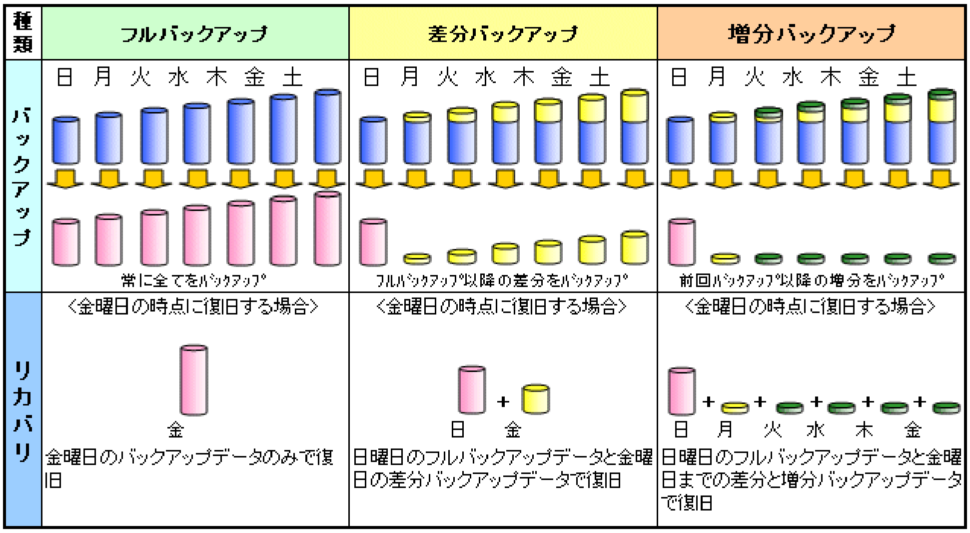


図 8 - 1：バックアップの実施方法

表 8 - 1：オフラインバックアップの方法

項目	バックアップ方法		
	フルバックアップ	差分バックアップ	増分バックアップ
バックアップサイズ	大きい	小さい	小さい
バックアップ時間	長い	～中	短い
フルリカバリー時間	短い	中～	長い
管理のしやすさ	容易	やや複雑	複雑

8.1.2 バックアップの構成

オフラインバックアップには、以下の 2 種類があります。

■ ローカルバックアップ／リストア

ローカルバックアップ／リストアでは、バックアップ／リストア対象サーバに搭載（または直結）したバックアップデバイスにバックアップデータを保存します。

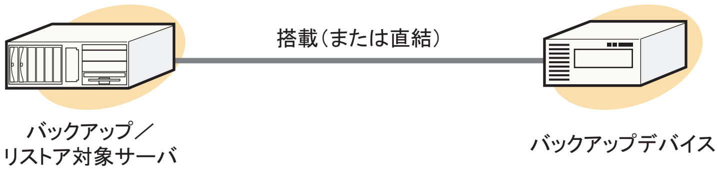


図 8 - 2：**ローカルバックアップ／リストアの構成**

■ リモートバックアップ／リストア

リモートバックアップ／リストアでは、バックアップデバイスを搭載（または直結）したサーバ（バックアップサーバ）に、ネットワークを経由してバックアップデータを保存します。

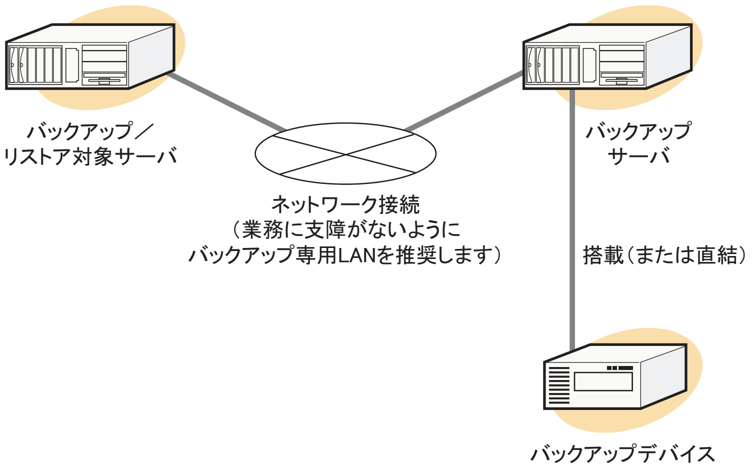


図 8 - 3：**リモートバックアップ／リストアの構成**



- **以下の環境では使用しないでください。**
  - **インターネットを経由した環境（ただし、**VPN **などを用いてセキュアな通信環境が確保されている場合は除外します。）**
  - NAT ／ NAPT **を利用した環境**
  - **ファイアーウォールやプロキシを経由した環境**

8.1.3　準備と環境確認

バックアップおよびリストアを行うために必要な準備と環境確認について説明します。
本書では、OS 標準コマンドによるオフラインバックアップについてのみ説明します。

参 考

より高度な機能を必要とされる場合

- **バックアップのスケジューリングやレポート機能およびグラフィカル画面での操作などといった、より高度な機能については、本書に記載しているコマンドでは対応できません。**
  より高度な機能を必要とされる場合は、NetVault を使用してバックアップ／リストアを行うことを推奨します。

1　**rpm コマンドを使用して、バックアップおよびリストアで使用するコマンドがインストールされているか確認します。**

インストールされていない場合は、バックアップするシステムにインストールしてください。
確認するコマンドは、以下のとおりです。

表 8 - 2：確認するコマンド

コマンド	パッケージ名	説明
mt	mt-st	テープの操作
dump	dump	バックアップに関する操作
restore	dump	リストアに関する操作

2　**バックアップおよびリストアを行うシステムに必要な情報を取得します。**

1. **"/etc/fstab" の内容を確認します。**

   /etc/fstab の例：

```
LABEL=/           /          ext3    defaults         1  1
LABEL=/boot       /boot      ext3    defaults         1  2
devpts            /dev/pts   devpts  gid=5,mode=620   0  0
tmpfs             /dev/shm   tmpfs   defaults         0  0
proc              /proc      proc    defaults         0  0
sysfs             /sys       sysfs   defaults         0  0
LABEL=SWAP-sda3   swap       swap    defaults         0  0
```

2. **以下のコマンドを実行し、パーティションごとの容量、順序、開始、および終了セクタ番号を確認します。**
   例：fdisk **コマンドによる確認（デバイスが** /dev/sda **で、パーティションテーブルが** msdos **形式の場合）**

```
# fdisk -l -u /dev/sda
Disk /dev/sda: 146.8 MB, 146815737856 bytes
255 heads, 63 sectors/track, 17849 cylinders, total 286749488 sectors
Units = セクタ数 of 1 * 512 = 512 bytes

デバイス Boot    Start       End     Blocks    Id   System
/dev/sda1 *         63    530144     265041    83   Linux
/dev/sda2       530145  42473185  41943041+    83   Linux
/dev/sda3     42473186  43525444   1052259+    82   Linux swap / Solaris
```

   例：parted **コマンドによる確認（パーティションテーブルが** gpt **形式の場合）**

```
# parted /dev/sda
(parted) unit s
(parted) print

モデル: LSI MegaRAID SAS RMB (scsi)
ディスク /dev/sda: 286749487s
セクタサイズ (論理/物理): 512B/512B
パーティションテーブル: gpt

番号  開始        終了        サイズ     タイプ    ファイルシステム  フラグ
 1   63s        530144s    530082s    プライマリ  ext3            boot
 2   530145s    42473185s  41943041s  プライマリ  ext3
 3   42473186s  43525444s  1052259s   プライマリ  linux-swap
(parted) quit
```

3. **以下のコマンドを実行し、ファイルシステムのラベル名を確認します。**

```
# e2label /dev/sda1
/boot
# e2label /dev/sda2
/
```

4. swap **のパーティションに設定されているラベル名を確認する場合は、以下のコマンドを実行し、**TYPE="swap" **のデバイスを表示させます。**

```
# blkid | grep swap
/dev/sda3: LABEL="SWAP-sda3" TYPE="swap"
```

8.2 バックアップ

オフラインバックアップの実施手順を説明します。

8.2.1 ランレベル 1（シングルユーザーモード）への移行

バックアップを行う前に、ファイルの整合性を確保するため、必要なプロセスを除くすべてのプロセスを停止してください。

注 意

- **各種サービスが完全に停止せず、バックアップ対象となるボリュームが使用中の状態になる場合があるため、システムをランレベル 1（シングルユーザーモード）に移行する際は、通常運用の状態からの init コマンドによる移行は行わないでください。**

1 **以下のコマンドを実行し、/boot/grub/grub.conf の編集を開始します。**

```
# vi /boot/grub/grub.conf
```

2 **ご使用のカーネル版数、ファイルシステムのラベル名をランレベル 1（シングルユーザーモード）の起動エントリーに追加します。**

```
default=0
timeout=5
splashimage=(hd0,0)/grub/splash.xpm.gz
hiddenmenu
title Red Hat Enterprise Linux Server (2.6.18-274.el5)
          root (hd0,0)
          kernel /vmlinuz-2.6.18-274.el5 ro root=LABEL=/ rhgb quiet mce=0
nmi_watchdog=0
          initrd /initrd-2.6.18-274.el5.img
title Red Hat Enterprise Linux Server (2.6.18-274.el5) Single User Mode
          root (hd0,0)
          kernel /vmlinuz-2.6.18-274.el5 ro root=LABEL=/ rhgb quiet 1
          initrd /initrd-2.6.18-274.el5.img
```

追加したエントリーの kernel 行の末尾に、下線のとおり「1」を追記します。これは、ランレベル 1（シングルユーザーモード）で起動することを示します。

3 **システムを再起動します。**

```
# shutdown -r now
```

4 **グラフィカル GRUB ブートローダ画面で追加したエントリーを選択し、ランレベル 1（シングルユーザーモード）に移行します。**

8.2.2 ファイルシステムの整合性チェック

ファイルシステムの整合性をチェックし、正常な状態かどうかを確認します。

■ アンマウント可能なファイルシステムの場合

ファイルシステムがアンマウント可能な場合、ファイルシステムをアンマウントした状態で実施します。

● ext3 ファイルシステムの場合

/boot にマウントしている /dev/sda1 ボリューム（ext3 ファイルシステム）をチェックする場合の手順を例に説明します。

1 **以下のコマンドを実行し、バックアップ対象のファイルシステムをアンマウントします。**

```
# cd /
# umount /boot
```

注 意

- **umount コマンドがエラーとなった場合、e2fsck コマンドを実行しないでください。エラーが発生する場合は、ファイルシステムが何らかの原因で使用中である可能性があります。原因を特定して、ファイルシステムが使用されていない状態にしてから、再度実行してください。**

2 以下のコマンドを実行し、アンマウントしたボリュームをチェックします。

```
# e2fsck -nfv /dev/sda1
```

ファイルシステムのエラーを検出した場合、ディストリビューションのブータブルメディアからレスキューモードの起動を行い、該当ボリュームのファイルシステムを復旧してください。レスキューモードの起動方法については、「7.4 レスキューモードの起動（トラブル対応時）」（→ P.154）を参照してください。

■ "/"（ルート）などアンマウント不可能なファイルシステムの場合

"/"（ルート）にマウントしている /dev/sda2 ボリュームをチェックする場合の手順を例に説明します。

1 以下のコマンドを実行し、"/"（ルート）ファイルシステムをリードオンリーボリュームとしてリマウントします。

```
# cd /
# mount -r -n -o remount /
```

注 意

- **リードオンリーでのマウントがエラーとなった場合、e2fsck コマンドを実行しないでください。ファイルシステムが マウント状態のまま e2fsck コマンドを実行すると、チェックしているボリュームが破壊される可能性があります。「8.2.1 ランレベル 1（シングルユーザーモード）への移行」（→ P.163）を参照して、ランレベル 1（シングルユーザーモード）でシステムを再起動し、再度コマンドを実行してください。**

2 以下のコマンドを実行し、リードオンリーでリマウントしたボリュームをチェックします。

```
# e2fsck -nfv /dev/sda2
```

ファイルシステムのエラーを検出した場合、ディストリビューションのブータブルメディアからレスキューモードの起動を行って、該当ボリュームのファイルシステムを復旧してください。レスキューモードの起動方法については、「7.4 レスキューモードの起動（トラブル対応時）」（→ P.154）を参照してください。

3 以下のコマンドを実行し、チェック済みのボリュームをもとのマウント状態に戻します。

```
# mount -w -n -o remount /
```

8.2.3 ファイルシステムのローカルバックアップ（テープ装置を使用の場合）

ディスクのパーティション単位に、ローカルに接続されたテープ装置（DAT、LTO）を使用してバックアップを行います。

注 意

- **テープ装置のブロックサイズは、バックアップソフトウェア、他者使用などによって変更されることがあります。そのため、テープ装置を使用する際は、ブロックサイズを設定してください。テープ装置が他のマシンと共有されている場合は、テープ装置が他のマシンで利用されていない状態でブロックサイズを設定してください。**
  例：可変長に設定する場合
  ```
  # mt -f /dev/nst0 setblk 0
  ```
  例：固定長 10240byte に設定する場合
  ```
  # mt -f /dev/nst0 setblk 10240
  ```
- **テープ装置を固定長ブロックで使用する場合、バックアップコマンドに同じブロックサイズを指定してください。dump コマンドの場合、[-b] オプションを使用します。また、リストア時には、同じブロックサイズを指定してください。**
- **ブロックサイズは 512byte の倍数で指定します。ブロックサイズは使用するテープ装置ごとに最大値が異なります。詳細は、各テープ装置のマニュアルを参照してください。**
- **テープ装置が他のマシンと共有されているなどの理由により、テープ装置がロックされ、利用できなくなる場合があります。その場合は、ロックされた原因を取り除いて利用可能な状態にしてください。**

参 考

- **mt コマンドや dump コマンドに指定するテープ装置（DAT、LTO）のデバイス名は、以下のように指定します。**
  - **読み込み／書き出し完了時、自動的にテープの巻き戻しを行いたい場合**
    /dev/st0
  - **読み込み／書き出し完了時、自動的なテープの巻き戻しを行いたくない場合**
    /dev/nst0

1 **テープのツメが書き込み可能になっていることを確認し、テープをテープデバイスに挿入します。**

テープデバイスを認識しない場合は、以下のコマンドを実行します。

```
# mknod /dev/nst0 c 9 128
```

2 **以下のコマンドを実行し、テープを巻き戻します。**

```
# mt -f /dev/nst0 rewind
```

3 以下のコマンドを実行してテープユニットの状態を表示し、ファイル番号（File number）を確認します。

```
# mt -f /dev/nst0 status
```

テープの先頭からバックアップデータを格納するため、最初はファイル番号（File number）が「0」であることを確認してください。

4 以下のコマンドを実行し、ファイルシステムパーティションをバックアップします。

例：ext3 ファイルシステムの場合

```
# dump 0uf /dev/nst0 /dev/sda1
```

「8.1.3 準備と環境確認」（→ P.161）の構成では、上記は "/boot" のパーティションになります。

注意

- **dump コマンドにブロックサイズを指定する場合は、[-b] オプションを使用してください。**

参考

dump コマンドの実行速度を上げるには

- **dump コマンドを実行する前に以下のコマンドを実行すると、dump コマンドの実行速度が上がります。**

```
# echo 'deadline' > /sys/block/sda/queue/scheduler
```

「8.1.3 準備と環境確認」（→ P.161）の構成では、デバイス名が sda なのでリダイレクト先が /sys/block/sda/queue/scheduler となります。バックアップ対象のデバイスに従って上記のファイル名を適切に指定してください。

5 以下のコマンドを実行してテープユニットの状態を表示し、ファイル番号（File number）を確認します。

mt コマンドの status を実行しても、テープの書き込み位置は変更されません。

```
# mt -f /dev/nst0 status
```

ファイル番号が、dump コマンドでバックアップする前よりも「1」増えていることを確認します。

参考

- **ファイル番号が増えていない場合は、以下のコマンドを実行して再度テープユニットの情報を確認してください。**

```
# mt -f /dev/nst0 asf <ファイル番号>
```

***<ファイル番号>*は、dump コマンドでバックアップする前のファイル番号に「1」加算した値です。**

注 意

- **ファイル番号とバックアップしたパーティション名は必ずメモしておいてください。**

6 以下のコマンドを実行し、ファイルシステムパーティションをバックアップします。

*<パーティションデバイス名>*に "/dev/sda2" を設定すると、「8.1.3 準備と環境確認」（→ P.161）の構成では "/"（ルート）のパーティションになります。

```
# dump 0uf /dev/nst0 <パーティションデバイス名>
```

注 意

- dump **コマンドにブロックサイズを指定する場合は、**[-b] **オプションを使用してください。**

7 すべてのファイルシステムパーティションをバックアップするまで、手順 5 ～手順 6 を繰り返します。

バックアップの途中でテープを使い切った場合は、新しいテープに入れ替えて、そのまま継続してください。

8 すべてのファイルシステムパーティションのバックアップが完了したら、以下のコマンドを実行し、テープを取り出します。

```
# mt -f /dev/nst0 offline
```

バックアップに使用したテープは、書き込み禁止の状態で保管してください。

9 以下のコマンドを実行しシステムを再起動して、通常運用状態に戻します。

```
# shutdown -r now
```

8.2.4 ファイルシステムのローカルバックアップ（データカートリッジを使用の場合）

データカートリッジを使用できる機種の場合、ディスクのパーティション単位に、ローカルに接続されたデータカートリッジを使用してバックアップを行うことができます。

注 意

- **データカートリッジを使用できるかどうかは、ご使用の機種の『システム構成図』で確認してください。**
- **1 つのパーティションを複数のデータカートリッジにまたがってバックアップすることはできません。また、データカートリッジの空き容量は、バックアップ対象となるパーティションの使用量よりも大きいことを必ず確認してください。**
- **バックアップデータの保存先となるデータカートリッジには、あらかじめラベル名を付与してください。**

1 以下のコマンドを実行し、データカートリッジにあるバックアップ先をマウントします。

マウントポイントを作成してデータカートリッジにあるバックアップ先をマウントします（すでにマウント済みの場合は、本手順は必要ありません)。

```
# mkdir /backup
# mount /dev/sdc1 /backup
```

参 考

- **ここでは、データカートリッジは /dev/sdc とし、パーティション 1（/dev/sdc1）にバックアップすると仮定しています。データカートリッジのデバイス名を確認する方法については、ご使用のデータカートリッジの取扱説明書を参照してください。**
  データカートリッジがフォーマットされていない場合は、以下のように ext3 でフォーマットしてからマウントしてください。

```
# mkfs.ext3 /dev/sdc1
# e2label /dev/sdc1 /backlabel
```

2 以下のコマンドを実行し、ファイルシステムパーティションをバックアップします。

例：ext3 ファイルシステムの場合

```
# dump 0uf /backup/backupdata1.dat /dev/sda1
```

「8.1.3 準備と環境確認」（→ P.161）の構成では、上記は "/boot" のパーティションになります。

参考

dump コマンドの実行速度を上げるには

- dump コマンドを実行する前に以下のコマンドを実行すると、dump コマンドの実行速度が上がります。

```
# echo 'deadline' > /sys/block/sda/queue/scheduler
```

「8.1.3 準備と環境確認」（→ P.161）の構成では、デバイス名が sda なのでリダイレクト先が /sys/block/sda/queue/scheduler となります。バックアップ対象のデバイスに従って上記のファイル名を適切に指定してください。

3 **次のファイルシステムパーティションをバックアップします。**

*<パーティションデバイス名>*に "/dev/sda2" を設定すると、「8.1.3 準備と環境確認」（→ P.161）の構成では "/"（ルート）のパーティションになります。

```
# dump 0uf /backup/<バックアップファイル名> <パーティションデバイス名>
```

4 **すべてのファイルシステムパーティションをバックアップするまで、手順 3 を繰り返します。**

5 **すべてのファイルシステムパーティションのバックアップが完了したら、アンマウントし、ディレクトリを削除してください。**

```
# umount /backup
# rm -rf /backup
```

6 **以下のコマンドを実行してシステムを再起動し、通常運用状態に戻します。**

```
# shutdown -r now
```

参考

- 複数のデータカートリッジを使用する場合は、データカートリッジ交換後、データカートリッジを再認識させるために「8.2.1 ランレベル 1（シングルユーザーモード）への移行」（→ P.163）と「8.2.4 ファイルシステムのローカルバックアップ（データカートリッジを使用の場合）」（→ P.169）の手順を繰り返してください。

8.2.5 ファイルシステムのリモートバックアップ（テープ装置を使用の場合）

ディスクのパーティション単位に、バックアップサーバのテープ装置（DAT、LTO）を使用してバックアップを行います。

注 意

- **ネットワークを使用可能な状態にしてください。ネットワークは、セキュリティを維持した環境および業務に支障がないようバックアップ専用 LAN を用意することを推奨します。**
- **テープ装置のブロックサイズは、バックアップソフトウェア、他者使用などによって変更されることがあります。そのため、テープ装置を使用する際は、ブロックサイズを設定してください。テープ装置が他のマシンと共有されている場合は、テープ装置が他のマシンで利用されていない状態でブロックサイズを設定してください。**
  例：可変長に設定する場合
  ```
  # ssh <IPアドレスまたはホスト名> mt -f /dev/nst0 setblk 0
  ```
  例：固定長 10240byte に設定する場合
  ```
  # ssh <IPアドレスまたはホスト名> mt -f /dev/nst0 setblk 10240
  ```
- **テープ装置を固定長ブロックで使用する場合、バックアップコマンドに同じブロックサイズを指定してください。dump コマンドでは [-b] オプション、dd コマンドでは [obs] オプションを使用します。また、リストア時には、同じブロックサイズを指定してください。**
- **ブロックサイズは 512byte の倍数で指定します。ブロックサイズは使用するテープ装置ごとに最大値が異なります。詳細は、各テープ装置のマニュアルを参照してください。**
- **テープ装置が他のマシンと共有されているなどの理由により、テープ装置がロックされ、利用できなくなる場合があります。その場合は、ロックされた原因を取り除いて利用可能な状態にしてください。**
- **バックアップサーバへの操作は ssh コマンドを使用します。パスワードの問い合わせがある場合は、パスワードを入力します。**
- **1 つのパーティションをバックアップする際、複数のテープにまたがってバックアップすることはできません。**

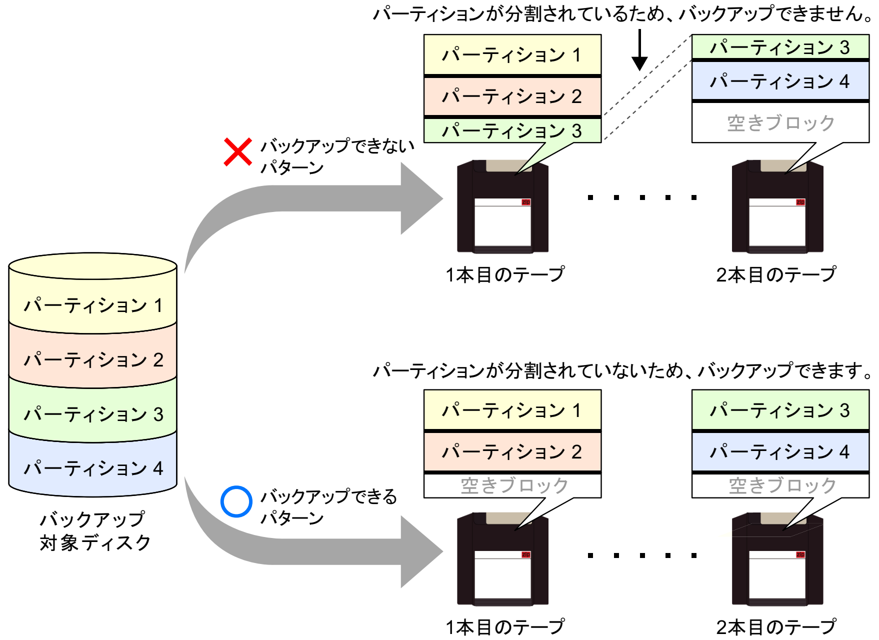


図 8 - 4：テープ装置を使用したリモートバックアップ

参 考

- mt コマンドや dump コマンドに指定するテープ装置（DAT、LTO）のデバイス名は、以下のように指定します。
  - 読み込み／書き出し完了時、自動的にテープの巻き戻しを行いたい場合
    /dev/st0
  - 読み込み／書き出し完了時、自動的なテープの巻き戻しを行いたくない場合
    /dev/nst0

1 以下のコマンドを実行し、ネットワークを起動します。

```
# /etc/rc.d/init.d/network start
```

2 テープのツメが書き込み可能になっていることを確認し、テープをテープデバイスに挿入します。

テープデバイスを認識しない場合は、以下のコマンドを実行します。

```
# ssh <IPアドレスまたはホスト名> mknod /dev/nst0 c 9 128
```

3 以下のコマンドを実行し、テープを巻き戻します。

```
# ssh <IPアドレスまたはホスト名> mt -f /dev/nst0 rewind
```

4 以下のコマンドを実行してテープユニットの状態を表示し、ファイル番号（File number）を確認します。

```
# ssh <IPアドレスまたはホスト名> mt -f /dev/nst0 status
```

テープの先頭からバックアップデータを格納するため、最初はファイル番号（File number）が「0」であることを確認してください。

5 以下のコマンドを実行し、ファイルシステムパーティションをバックアップします。

例：ext3 ファイルシステムの場合

```
# dump 0uf - /dev/sda1 | ssh <IPアドレスまたはホスト名> dd of=/dev/nst0
```

「8.1.3 準備と環境確認」（→ P.161）の構成では、上記は "/boot" のパーティションになります。

注 意

- **dd コマンドにブロックサイズを指定する場合は、[obs] オプションを使用してください。**

参 考

dump コマンドの実行速度を上げるには

- **dump コマンドを実行する前に以下のコマンドを実行すると、dump コマンドの実行速度が上がります。**

```
# echo 'deadline' > /sys/block/sda/queue/scheduler
```

「8.1.3 準備と環境確認」（→ P.161）の構成では、デバイス名が sda なのでリダイレクト先が /sys/block/sda/queue/scheduler となります。バックアップ対象のデバイスに従って上記のファイル名を適切に指定してください。

6 以下のコマンドを実行してテープユニットの状態を表示し、ファイル番号（File number）を確認します。

mt コマンドの status を実行しても、テープの書き込み位置は変更されません。

```
# ssh <IPアドレスまたはホスト名> mt -f /dev/nst0 status
```

ファイル番号が、dump コマンドでバックアップする前よりも「1」増えていることを確認します。

参 考

- **ファイル番号が増えていない場合は、以下のコマンドを実行して再度テープユニットの情報を確認してください。**

```
# ssh <IPアドレスまたはホスト名> mt -f /dev/nst0 asf <ファイル番号>
```

***<ファイル番号>*は、dump コマンドでバックアップする前のファイル番号に「1」加算した値です。**

注意

- **ファイル番号とバックアップしたパーティション名は必ずメモしておいてください。**

7 以下のコマンドを実行し、ファイルシステムパーティションをバックアップします。

*＜パーティションデバイス名＞*に "/dev/sda2" を設定すると、「8.1.3 準備と環境確認」（→ P.161）の構成では "/"（ルート）のパーティションになります。

```
# dump 0uf - <パーティションデバイス名> | ssh ¥
<IPアドレスまたはホスト名> dd of=/dev/nst0
```

注意

- **dd コマンドにブロックサイズを指定する場合は、[obs] オプションを使用してください。**

8 すべてのファイルシステムパーティションをバックアップするまで、手順 6 〜 手順 7 を繰り返します。

9 すべてのファイルシステムパーティションのバックアップが完了したら、以下のコマンドを実行し、テープを取り出します。

```
# ssh <IPアドレスまたはホスト名> mt -f /dev/nst0 offline
```

バックアップに使用したテープは、書き込み禁止の状態で保管してください。

10 以下のコマンドを実行しシステムを再起動して、通常運用状態に戻します。

```
# shutdown -r now
```

8.2.6 ファイルシステムのリモートバックアップ（データカートリッジを使用の場合）

バックアップサーバがデータカートリッジを使用できる機種の場合、ディスクのパーティション単位に、バックアップサーバのデータカートリッジを使用してバックアップを行うことができます。

注 意

- **データカートリッジを使用できるかどうかは、ご使用の機種の『システム構成図』で確認してください。**
- **1 つのパーティションを複数のデータカートリッジにまたがってバックアップすることはできません。また、データカートリッジの空き容量は、バックアップ対象となるパーティションの使用量よりも大きいことを必ず確認してください。**
- **バックアップデータの保存先となるデータカートリッジには、あらかじめラベル名を付与してください。**
- **ネットワークを使用可能な状態にしてください。ネットワークは、セキュリティを維持した環境および業務に支障がないようバックアップ専用 LAN を用意することを推奨します。**
- **バックアップサーバへの操作は ssh コマンドを使用します。パスワードの問い合わせがある場合は、パスワードを入力します。**

1 以下のコマンドを実行し、ネットワークを起動します。

```
# /etc/rc.d/init.d/network start
```

2 以下のコマンドを実行し、データカートリッジにあるバックアップ先をマウントします。

マウントポイントを作成してデータカートリッジにあるバックアップ先をマウントします（すでにマウント済みの場合は、本手順は必要ありません）。

```
# ssh <IPアドレスまたはホスト名> mkdir /backup
# ssh <IPアドレスまたはホスト名> mount /dev/sdc1 /backup
```

参 考

- **ここでは、データカートリッジは /dev/sdc とし、パーティション 1（/dev/sdc1）にバックアップすると仮定しています。データカートリッジのデバイス名を確認する方法については、ご使用のデータカートリッジの取扱説明書を参照してください。**
  データカートリッジがフォーマットされていない場合は、以下のように ext3 でフォーマットしてからマウントしてください。

```
# ssh <IPアドレスまたはホスト名> mkfs.ext3 /dev/sdc1
# ssh <IPアドレスまたはホスト名> e2label /dev/sdc1 /backlabel
```

3　以下のコマンドを実行し、ファイルシステムパーティションをバックアップします。

例：ext3 ファイルシステムの場合

```
# dump 0f - /dev/sda1 | ssh <IPアドレスまたはホスト名> ¥
dd of=/backup/backupdata1.dat
```

「8.1.3 準備と環境確認」（→ P.161）の構成では、上記は "/boot" のパーティションになります。

参　考

dump コマンドの実行速度を上げるには

- **dump コマンドを実行する前に以下のコマンドを実行すると、dump コマンドの実行速度が上がります。**

```
# echo 'deadline' > /sys/block/sda/queue/scheduler
```

「8.1.3 準備と環境確認」（→ P.161）の構成では、デバイス名が sda なのでリダイレクト先が /sys/block/sda/queue/scheduler となります。バックアップ対象のデバイスに従って上記のファイル名を適切に指定してください。

4　次のファイルシステムパーティションをバックアップします。

*＜パーティションデバイス名＞*に "/dev/sda2" を設定すると、「8.1.3 準備と環境確認」（→ P.161）の構成では "/"（ルート）のパーティションになります。

```
# dump 0f - <パーティションデバイス名> | ssh <IPアドレスまたはホスト名> ¥
dd of=/backup/<バックアップファイル名>
```

5　すべてのファイルシステムパーティションをバックアップするまで、手順 4 を繰り返します。

6　すべてのファイルシステムパーティションのバックアップが完了したら、アンマウントし、ディレクトリを削除してください。

```
# ssh <IPアドレスまたはホスト名> umount /backup
# ssh <IPアドレスまたはホスト名> rm -rf /backup
```

7　以下のコマンドを実行してシステムを再起動し、通常運用状態に戻します。

```
# shutdown -r now
```

8.3 リストア

リストアの方法について、実施手順を説明します。

注 意

- **ディスク故障の場合には、リストアを実施する前に故障したディスクを交換し、アレイ構成を設定してください。詳しくは、『運用管理マニュアル』または各ハードウェアの交換マニュアルを参照してください。**
- **システムボリュームをリストアする場合は、レスキューモードでシステムを起動してからリストアする必要があります（「■ システムボリュームをリストアする場合」（→ P.177））。**
- **システムボリューム以外をリストアする場合は、システムをランレベル 1（シングルユーザーモード）で起動してからリストアを実施します（「■ システムボリューム以外をリストアする場合」（→ P.178））。また、システムをリストアする場合と同様にレスキューモードにしてからリストアすることも可能です。**

8.3.1 システムの起動

リストアする場合は、リストア対象によって起動方法が変わります。リストア対象に応じて起動方法を適宜選択し、実施してください。

■ システムボリュームをリストアする場合

システムボリュームをリストアする場合は、インストール CD を用いて、レスキューモードでシステムを再起動してからリストアする必要があります。
レスキューモードの起動方法については、「7.4 レスキューモードの起動（トラブル対応時）」（→ P.154）を参照してください。

■ システムボリューム以外をリストアする場合

システムボリューム以外をリストアする場合は、ランレベル 1（シングルユーザーモード）にしてからリストアします。移行方法については、「8.2.1 ランレベル 1（シングルユーザーモード）への移行」（→ P.163）を参照してください。
なお、システムをリストアする場合と同様に、レスキューモードでシステム以外をリストアする場合は、「7.4 レスキューモードの起動（トラブル対応時）」（→ P.154）を参照してください。

注 意

- **各種サービスが完全に停止せず、リストア対象となるボリュームが使用中の状態になる場合があるため、システムをランレベル 1（シングルユーザーモード）に移行する際は、通常運用の状態からの init コマンドによる移行は行わないでください。**
  また、安全のため、ランレベル 1（シングルユーザーモード）へ移行したときは、リストア対象となる資源が使用中になっていないことを必ず確認してください。

8.3.2 パーティションの再構築（ディスク障害の場合）

parted コマンドまたは fdisk コマンドを使用して、ディスク上の領域をパーティションに分割します。
故障前と同じ環境にするためには、「8.1.3 準備と環境確認」（→ P.161）で確認したパーティションの開始セクタ番号、および終了セクタ番号が必要になります。

■ デバイスが /dev/sda で、パーティションテーブルが msdos 形式の場合

```
# fdisk /dev/sda
  主なコマンド:
    d:パーティションの削除
    p:パーティションの一覧
    n:パーティションの作成
    m:オンラインヘルプ
    q:作業結果を反映させずに終了
    w:作業結果を保存して終了
```

注 意

- **パーティションの分割が終わったら、必ずレスキューモードを exit コマンドで一度終了し、再度レスキューモードで起動してください。**
- **リストア対象のパーティションがシステムボリュームの場合は、ディスクラベルを "msdos" としてください。システムボリューム以外の場合は、使用環境に合わせてディスクラベルを設定してください。**
- **ディスクラベルを設定する場合は、parted コマンドを使用してください。**

■ パーティションテーブルが gpt 形式の場合

```
# parted /dev/sda
  主なコマンド:
    mklabel :ディスクラベルを設定
    rm      :パーティションの削除
    print   :パーティションの一覧
    mkpart  :パーティションの作成
    help    :オンラインヘルプ
    quit    :作業結果を保存して終了
```

注 意

- **パーティションの分割が終わったら、必ずレスキューモードを reboot コマンドで一度終了し、再度レスキューモードで起動してください。**
- **リストア対象のパーティションがシステムボリュームの場合は、ディスクラベルを "gpt" としてください。システムボリューム以外の場合は、使用環境に合わせてディスクラベルを設定してください。**

8.3.3 パーティション（ext3、swap）のフォーマット

ディスク上のパーティション（ext3、swap）を初期化します。

■ ext3 のパーティションの場合

1 **ext3 でフォーマットします。**

```
# mkfs.ext3 /dev/sda1
# mkfs.ext3 /dev/sda2
```

参 考

- **ext2 フォーマットを行う場合は、mkfs.ext2 コマンドを使用してください。**

2 **「8.1.3 準備と環境確認」（→ P.161）****で調べておいた、ファイルシステムのラベル名を設定します。**

```
# e2label /dev/sda1 /boot
# e2label /dev/sda2 /
```

swap のパーティションに「8.1.3 準備と環境確認」（→ P.161）で確認したラベル名を設定する場合は、「■ swap のパーティションの場合」（→ P.180）を参照してください。

3　ext3 ファイルシステムの［dir_index］オプションを無効にします。

1. **以下のコマンドを実行し、各デバイスにおいて、「Filesystem features:」の項目に「dir_index」があることを確認します。**

```
# tune2fs -l /dev/sda2 | grep features
Filesystem features: has_journal ext_attr resize_inode dir_index
filetype needs_recovery sparse_super …
```

2. **以下のコマンドを実行し、「dir_index」を無効にします。**

```
# tune2fs -O ^dir_index /dev/sda2
```

3. **以下のコマンドを実行し、「dir_index」がないことを確認します。**

```
# tune2fs -l /dev/sda2 | grep features
Filesystem features: has_journal ext_attr resize_inode filetype
needs_recovery sparse_super …
```

■ swap のパーティションの場合

以下のコマンドを実行し、swap パーティションをフォーマットします。なお、「8.1.3 準備と環境確認」（→ P.161）では、swap パーティションにラベル名「SWAP-sda3」が設定されているため、必ず [-L] オプションでラベル名を指定してください。

```
# /mnt/runtime/usr/sbin/mkswap -L SWAP-sda3 /dev/sda3
```

参　考

- 「8.1.3 準備と環境確認」（→ P.161）**で swap パーティションにラベル名が設定されていない場合は、[-L] オプションでラベル名を指定する必要はありません。**
  以下のように swap パーティションをフォーマットしてください。

```
# /mnt/runtime/usr/sbin/mkswap /dev/sda3
```

8.3.4　ファイルシステムのローカルリストア（テープ装置を使用の場合）

ディスクのパーティション単位に、ローカルに接続されたテープ装置（DAT、LTO）を使用してリストアを行います。

注　意

- **テープ装置のブロックサイズは、バックアップソフトウェア、他者使用などによって変更されることがあります。そのため、テープ装置を使用する際は、ブロックサイズを設定してください。テープ装置が他のマシンと共有されている場合は、テープ装置が他のマシンで利用されていない状態でブロックサイズを設定してください。**
  1. **nst0 のデバイスを以下のコマンドを実行して作成します。**

```
# mknod /tmp/nst0 c 9 128
```

下線部分は、「c」はキャラクターデバイス、「9」はテープ装置、「128」は装置番号を意味します。

2. 他のバックアップソフトを停止後、mt コマンドでブロックサイズを変更します。
   例：可変長に設定する場合

```
# mt -f /tmp/nst0 setblk 0
```

   例：固定長 10240byte に設定する場合

```
# mt -f /tmp/nst0 setblk 10240
```

- バックアップ時に実行したコマンドでブロックサイズを指定した場合は、リストア時にも同じブロックサイズを指定してください。restore コマンドの場合、[-b] オプションを使用します。
- テープ装置が他のマシンと共有されているなどの理由により、テープ装置がロックされ、利用できなくなる場合があります。その場合は、ロックされた原因を取り除いて利用可能な状態にしてください。

参 考

- mt コマンドや restore コマンドに指定するテープ装置（DAT、LTO）のデバイス名は、以下のように指定します。
  - 読み込み／書き出し完了時、自動的にテープの巻き戻しを行いたい場合
    /tmp/st0
  - 読み込み／書き出し完了時、自動的なテープの巻き戻しを行いたくない場合
    /tmp/nst0

1 テープのツメが書き込み禁止になっていることを確認し、テープをテープデバイスに挿入します。

2 テープデバイスを認識しない場合は、以下のコマンドを実行し、nst0 デバイスを作成します。

```
# mknod /tmp/nst0 c 9 128
```

3 以下のコマンドを実行し、作業用ディレクトリを作成します。

```
# mkdir /mnt/work
```

4 以下のコマンドを実行し、テープを巻き戻します。

```
# mt -f /tmp/nst0 rewind
```

5 以下のコマンドを実行してテープユニットの状態を表示し、ファイル番号を確認します。

```
# mt -f /tmp/nst0 status
```

テープの先頭からリストアする場合は、ファイル番号が「0」であることを確認してください。

6 以下のコマンドを実行し、リストアするパーティションをマウントします。

例：ext3 ファイルシステムの場合

```
# mount -t ext3 <リストアするパーティションデバイス名> /mnt/work
```

注 意

- **「8.2.3 ファイルシステムのローカルバックアップ（テープ装置を使用の場合）」（→ P.166）でメモしたテープ内のファイル番号と、リストア先のパーティションデバイス名の関係を間違えないでください。**

7 以下のコマンドを実行し、ファイルシステムをマウントしたディレクトリに移動して、ファイルシステムをリストアします。

例：ext3 ファイルシステムの場合

```
# cd /mnt/work
# restore rf /tmp/nst0
```

本手順を実行後、「restore:/lost+found:File exists」のメッセージが出力される場合がありますが、問題ありません。

注 意

- **restore コマンドにブロックサイズを指定する場合は、[-b] オプションを使用してください。**
- **restore コマンドを実行すると、restoresymtable というファイルがカレントディレクトリに作成されます。すべてのリストアが終了するまでは削除しないでください。「8.3.8 ブートローダの設定（システムボリュームのリストア時）」（→ P.192）が終了し、システムが起動したあと、restoresymtable ファイルを削除してください。**
- **/tmp を独立したパーティションにしている場合、バックアップ／リストアを行った際に /tmp のパーミッションが 1777 から変更されてしまい、/tmp を使用する特定のサービスが起動できなくなる可能性があります。リストア実施後に、/tmp のパーミッションを 1777 に戻してください。/tmp ディレクトリのパーミッション変更手順は、以下のとおりです。**

```
# ls -l /mnt/work | grep tmp
dr--r--r-- 109 root root 8192 11月 19 04:02 tmp
# chmod 1777 /mnt/work/tmp
drwxrwxrwt 109 root root 8192 11月 19 04:02 tmp
```

8 以下のコマンドを実行し、リストアしたファイルシステムをアンマウントし、ファイルシステムをチェックします。

例：ext3 ファイルシステムの場合

```
# cd /
# umount /mnt/work
# e2fsck -nfv <リストアしたパーティションデバイス名>
# e2fsck -pfv <リストアしたパーティションデバイス名>
```

*<リストアしたパーティションデバイス名>*には、リストアした "/dev/sda1" などのパーティションデバイス名を設定してください。

注 意

- **e2fsck コマンドは、[-n] オプションでファイルシステムのチェックを行ったあと、データ削除を伴わない修復のみを行う [-p] オプションを使用してください。e2fsck コマンドの [-y] オプションを使用したチェックでは、ファイルシステムの整合性を保つために、データ削除を伴う修復も無条件で実施されます。この動作を意識せずに [-y] オプションを使用すると、意図しないデータ削除が行われる場合があるため、[-y] オプションは非推奨としています。**

9 **リストアするパーティションが残っている場合は、以下のコマンドを実行してテープユニットの状態を表示し、ファイル番号を確認します。**

mt コマンドの status を実行しても、テープの書き込み位置は変更されません。

```
# mt -f /tmp/nst0 status
```

ファイル番号が、restore コマンドでリストアする前よりも「1」増えていることを確認します。

参 考

- **増えていない場合は、以下のコマンドを実行して再度テープユニットの情報を確認してください。**

```
# mt -f /tmp/nst0 asf <ファイル番号>
```

***<ファイル番号>*は、restore コマンドでリストアする前のファイル番号に「1」加算した値です。**

10 **すべてのファイルシステムパーティションをリストアするまで、手順 6 ～手順 9 を繰り返します。**

複数のテープを使用したバックアップデータの場合は、画面の指示に従ってテープを入れ替えてください。

11 **すべてのパーティションのリストアが完了したら、以下のコマンドを実行してテープを取り出します。**

```
# mt -f /tmp/nst0 offline
```

12 **以下のコマンドを実行し、システムを再起動します。**

```
# reboot
```

13 **システムボリュームをリストアした場合は、続けて 「8.3.8 ブートローダの設定（システムボリュームのリストア時）」（→ P.192）を実施してください。**

8.3.5 ファイルシステムのローカルリストア（データカートリッジを使用の場合）

ディスクのパーティション単位に、ローカルに接続されたデータカートリッジを使用してリストアを行います。

1 **バックアップデータが格納されているデータカートリッジをマウントします。**

以下のコマンドを実行してマウントポイントを作成し、データカートリッジにあるバックアップ先をマウントします。

```
# mkdir /backup
# mount /dev/sdc1 /backup
```

2 **以下のコマンドを実行し、ラベル名を確認します。**

手順 1 でマウントしたパーティションが、バックアップ保存先として使用したパーティションであることをラベル名で確認します。

```
# e2label /dev/sdc1
/backlabel
```

3 **以下のコマンドを実行し、作業用ディレクトリを作成します。**

```
# mkdir /mnt/work
```

4 **以下のコマンドを実行し、リストアするパーティションをマウントします。**

例：ext3 **ファイルシステムの場合**

```
# mount -t ext3 <リストアするパーティションデバイス名> /mnt/work
```

5 **以下のコマンドを実行し、ファイルシステムをマウントしたディレクトリに移動して、ファイルシステムをリストアします。**

例：ext3 **ファイルシステムの場合**

```
# cd /mnt/work
# restore rf /backup/<バックアップファイル名>
```

本手順を実行後、「restore:/lost+found:File exists」のメッセージが出力される場合がありますが、問題ありません。

注 意

- restore コマンドを実行すると、restoresymtable というファイルがカレントディレクトリに作成されます。すべてのリストアが終了するまでは削除しないでください。「8.3.8 ブートローダの設定（システムボリュームのリストア時）」（→ P.192）が終了し、システムが起動したあと、restoresymtable ファイルを削除してください。
- /tmp を独立したパーティションにしている場合､バックアップ／リストアを行った際に /tmp のパーミッションが 1777 から変更されてしまい、/tmp を使用する特定のサービスが起動できなくなる可能性があります。リストア実施後に、/tmp のパーミッションを 1777 に戻してください。/tmp ディレクトリのパーミッション変更手順は、以下のとおりです。

```
# ls -l /mnt/work | grep tmp
dr--r--r-- 109 root root 8192 11月 19 04:02 tmp
# chmod 1777 /mnt/work/tmp
drwxrwxrwt 109 root root 8192 11月 19 04:02 tmp
```

6 以下のコマンドを実行してリストアしたファイルシステムをアンマウントし、ファイルシステムをチェックします。

例：ext3 ファイルシステムの場合

```
# cd /
# umount /mnt/work
# e2fsck -nfv <リストアしたパーティションデバイス名>
# e2fsck -pfv <リストアしたパーティションデバイス名>
```

*＜リストアしたパーティションデバイス名＞*には、リストアした "/dev/sda1" などのパーティションデバイス名を設定してください。

注 意

- e2fsck コマンドは、[-n] オプションでファイルシステムのチェックを行ったあと、データ削除を伴わない修復のみを行う [-p] オプションを使用してください。e2fsck コマンドの [-y] オプションを使用したチェックでは、ファイルシステムの整合性を保つために、データ削除を伴う修復も無条件で実施されます。この動作を意識せずに [-y] オプションを使用すると、意図しないデータ削除が行われる場合があるため、[-y] オプションは非推奨としています。

7 すべてのファイルシステムパーティションをリストアするまで、手順 4 〜手順 6 を繰り返します。

8 すべてのパーティションのリストアが完了したら、アンマウントしディレクトリを削除します。

```
# umount /backup
# rm -rf /backup
```

9 以下のコマンドを実行し、システムを再起動します。

```
# reboot
```

参 考

- 複数のデータカートリッジを使用する場合は、データカートリッジの交換後、データカートリッジを再認識させるために、レスキューモードの起動と「8.3.5 ファイルシステムのローカルリストア（データカートリッジを使用の場合）」（→ P.184）の手順を繰り返してください。
- レスキューモードの起動方法については、「7.4 レスキューモードの起動（トラブル対応時）」（→ P.154）を参照してください。

10 システムボリュームをリストアした場合は、続けて「8.3.8 ブートローダの設定（システムボリュームのリストア時）」（→ P.192）を実施してください。

8.3.6 ファイルシステムのリモートリストア（テープ装置を使用の場合）

ディスクのパーティション単位に、バックアップサーバのテープ装置（DAT、LTO）を使用してリストアを行います。

注 意

- ネットワークを使用可能な状態にしてください。ネットワークは、セキュリティを維持した環境および業務に支障がないようバックアップ専用 LAN を用意することを推奨します。
- テープ装置のブロックサイズは、バックアップソフトウェア、他者使用などによって変更されることがあります。そのため、テープ装置を使用する際は、ブロックサイズを設定してください。テープ装置が他のマシンと共有されている場合は、テープ装置が他のマシンで利用されていない状態でブロックサイズを設定してください。
  1. nst0 のデバイスを以下のコマンドを実行して作成します。
     ```
     # ssh <IPアドレスまたはホスト名> mknod /dev/nst0 c 9 128
     ```
     下線部分は、「c」はキャラクターデバイス、「9」はテープ装置、「128」は装置番号を意味します。
  2. 他のバックアップソフトウェアを停止後、mt コマンドでブロックサイズを変更します。
     例：可変長に設定する場合
     ```
     # ssh <IPアドレスまたはホスト名> mt -f /dev/nst0 setblk 0
     ```
     例：固定長 10240byte に設定する場合
     ```
     # ssh <IPアドレスまたはホスト名> mt -f /dev/nst0 setblk 10240
     ```
- バックアップ時に実行したコマンドでブロックサイズを指定した場合は、リストア時にも同じブロックサイズを指定してください。dd コマンドの場合、[ibs] オプションを使用します。
- テープ装置が他のマシンと共有されているなどの理由により、テープ装置がロックされ、利用できなくなる場合があります。その場合は、ロックされた原因を取り除いて利用可能な状態にしてください。
- バックアップサーバへの操作は ssh コマンドを使用します。パスワードの問い合わせがある場合は、パスワードを入力します。

参考

- mt コマンドや restore コマンドに指定するテープ装置（DAT、LTO）のデバイス名は、以下のように指定します。
  - 読み込み／書き出し完了時、自動的にテープの巻き戻しを行いたい場合
    /dev/st0
  - 読み込み／書き出し完了時、自動的なテープの巻き戻しを行いたくない場合
    /dev/nst0

1 **以下のコマンドを実行し、ネットワークを起動します。**

```
# /etc/rc.d/init.d/network start
```

注意

- レスキューモードで起動した場合は、ネットワークの起動は不要です。

2 **テープのツメが書き込み禁止になっていることを確認し、テープをテープデバイスに挿入します。**

テープデバイスを認識しない場合は、以下のコマンドを実行します。

```
# ssh <IPアドレスまたはホスト名> mknod /dev/nst0 c 9 128
```

3 **以下のコマンドを実行し、作業用ディレクトリを作成します。**

```
# mkdir /mnt/work
```

4 **以下のコマンドを実行し、テープを巻き戻します。**

```
# ssh <IPアドレスまたはホスト名> mt -f /dev/nst0 rewind
```

5 **以下のコマンドを実行してテープユニットの状態を表示し、ファイル番号を確認します。**

```
# ssh <IPアドレスまたはホスト名> mt -f /dev/nst0 status
```

テープの先頭からリストアする場合は、ファイル番号が「0」であることを確認してください。

6 **以下のコマンドを実行し、リストアするパーティションをマウントします。**

例：ext3 ファイルシステムの場合

```
# mount -t ext3 <リストアするパーティションデバイス名> /mnt/work
```

注意

- 「8.2.5 ファイルシステムのリモートバックアップ（テープ装置を使用の場合）」（→ P.171）でメモしたテープ内のファイル番号と、リストア先のパーティションデバイス名の関係を間違えないでください。

7 以下のコマンドを実行し、ファイルシステムをマウントしたディレクトリに移動して、ファイルシステムをリストアします。

例：ext3 ファイルシステムの場合

```
# cd /mnt/work
# ssh <IPアドレスまたはホスト名> dd if=/dev/nst0 | restore rf -
```

本手順を実行後、「restore:/lost+found:File exists」のメッセージが出力される場合がありますが、問題ありません。

注 意

- **dd コマンドにブロックサイズを指定する場合は、[ibs] オプションを使用してください。**
- **restore コマンドを実行すると、restoresymtable というファイルがカレントディレクトリに作成されます。すべてのリストアが終了するまでは削除しないでください。**「8.3.8 ブートローダの設定（システムボリュームのリストア時）」（→ P.192）**が終了し、システムが起動したあと、restoresymtable ファイルを削除してください。**
- **/tmp を独立したパーティションにしている場合、バックアップ／リストアを行った際に /tmp のパーミッションが 1777 から変更されてしまい、/tmp を使用する特定のサービスが起動できなくなる可能性があります。リストア実施後に、/tmp のパーミッションを 1777 に戻してください。/tmp ディレクトリのパーミッション変更手順は、以下のとおりです。**

```
# ls -l /mnt/work | grep tmp
dr--r--r-- 109 root root 8192 11月 19 04:02 tmp
# chmod 1777 /mnt/work/tmp
drwxrwxrwt 109 root root 8192 11月 19 04:02 tmp
```

8 以下のコマンドを実行し、リストアしたファイルシステムをアンマウントし、ファイルシステムをチェックします。

例：ext3 ファイルシステムの場合

```
# cd /
# umount /mnt/work
# e2fsck -nfv <リストアしたパーティションデバイス名>
# e2fsck -pfv <リストアしたパーティションデバイス名>
```

*<リストアしたパーティションデバイス名>*には、リストアした "/dev/sda1" などのパーティションデバイス名を設定してください。

注 意

- **e2fsck コマンドは、[-n] オプションでファイルシステムのチェックを行ったあと、データ削除を伴わない修復のみを行う [-p] オプションを使用してください。e2fsck コマンドの [-y] オプションを使用したチェックでは、ファイルシステムの整合性を保つために、データ削除を伴う修復も無条件で実施されます。この動作を意識せずに [-y] オプションを使用すると、意図しないデータ削除が行われる場合があるため、[-y] オプションは非推奨としています。**

9　リストアするパーティションが残っている場合は、以下のコマンドを実行してテープユニットの状態を表示し、ファイル番号を確認します。

mt コマンドの status を実行しても、テープの書き込み位置は変更されません。

```
# ssh <IPアドレスまたはホスト名> mt -f /dev/nst0 status
```

ファイル番号が、restore コマンドでリストアする前よりも「1」増えていることを確認します。

参 考

- **増えていない場合は、以下のコマンドを実行して再度テープユニットの情報を確認してください。**

  ```
  # ssh <IPアドレスまたはホスト名> mt -f /dev/nst0 asf <ファイル番号>
  ```

  ***<ファイル番号>*は、restore コマンドでリストアする前のファイル番号に「1」加算した値です。**

10　すべてのファイルシステムパーティションをリストアするまで、手順 6 〜手順 9 を繰り返します。

11　すべてのパーティションのリストアが完了したら、以下のコマンドを実行してテープを取り出します。

```
# ssh <IPアドレスまたはホスト名> mt -f /dev/nst0 offline
```

12　以下のコマンドを実行し、システムを再起動します。

```
# reboot
```

13　システムボリュームをリストアした場合は、続けて「8.3.8 ブートローダの設定（システムボリュームのリストア時）」（→ P.192）を実施してください。

8.3.7　ファイルシステムのリモートリストア（データカートリッジを使用の場合）

ディスクのパーティション単位に、ローカルに接続されたデータカートリッジを使用してリストアを行います。

注 意

- **ネットワークを使用可能な状態にしてください。ネットワークは、セキュリティを維持した環境および業務に支障がないようバックアップ専用 LAN を用意することを推奨します。**
- **バックアップサーバへの操作は ssh コマンドを使用します。パスワード入力画面が表示されたら、パスワードを入力します。**

1 以下のコマンドを実行し、ネットワークを起動します。

```
# /etc/rc.d/init.d/network start
```

注 意

- **レスキューモードで起動した場合は、ネットワークの起動は不要です。**

2 バックアップデータが格納されているデータカートリッジをマウントします。

以下のコマンドを実行してマウントポイントを作成し、データカートリッジにあるバックアップ先をマウントします。

```
# ssh <IPアドレスまたはホスト名> mkdir /backup
# ssh <IPアドレスまたはホスト名> mount /dev/sdc1 /backup
```

3 以下のコマンドを実行し、ラベル名を確認します。

手順 2 でマウントしたパーティションが、バックアップ保存先として使用したパーティションであることをラベル名で確認します。

```
# ssh <IPアドレスまたはホスト名> e2label /dev/sdc1
/backlabel
```

4 以下のコマンドを実行し、作業用ディレクトリを作成します。

```
# mkdir /mnt/work
```

5 以下のコマンドを実行し、リストアするパーティションをマウントします。

例：ext3 ファイルシステムの場合

```
# mount -t ext3 <リストアするパーティションデバイス名> /mnt/work
```

6 以下のコマンドを実行し、ファイルシステムをマウントしたディレクトリに移動して、ファイルシステムをリストアします。

例：ext3 ファイルシステムの場合

```
# cd /mnt/work
# ssh <IPアドレスまたはホスト名> dd if=/backup/<バックアップファイル名> |
restore rf -
```

本手順を実行後、「restore:/lost+found:File exists」のメッセージが出力される場合がありますが、問題ありません。

注 意

- restore コマンドを実行すると、restoresymtable というファイルがカレントディレクトリに作成されます。すべてのリストアが終了するまでは削除しないでください。システムが起動したあと、restoresymtable ファイルを削除してください。
- /tmp を独立したパーティションにしている場合、バックアップ／リストアを行った際に /tmp のパーミッションが 1777 から変更されてしまい、/tmp を使用する特定のサービスが起動できなくなる可能性があります。リストア実施後に、/tmp のパーミッションを 1777 に戻してください。/tmp ディレクトリのパーミッション変更手順は、以下のとおりです。

```
# ls -l /mnt/work | grep tmp
dr--r--r-- 109 root root 8192 11月 19 04:02 tmp
# chmod 1777 /mnt/work/tmp
drwxrwxrwt 109 root root 8192 11月 19 04:02 tmp
```

7 以下のコマンドを実行してリストアしたファイルシステムをアンマウントし、ファイルシステムをチェックします。

例：ext3 ファイルシステムの場合

```
# cd /
# umount /mnt/work
# e2fsck -nfv ＜リストアしたパーティションデバイス名＞
# e2fsck -pfv ＜リストアしたパーティションデバイス名＞
```

*＜リストアしたパーティションデバイス名＞*には、リストアした "/dev/sda1" などのパーティションデバイス名を設定してください。

注 意

- e2fsck コマンドは、[-n] オプションでファイルシステムのチェックを行ったあと、データ削除を伴わない修復のみを行う [-p] オプションを使用してください。e2fsck コマンドの [-y] オプションを使用したチェックでは、ファイルシステムの整合性を保つために、データ削除を伴う修復も無条件で実施されます。この動作を意識せずに [-y] オプションを使用すると、意図しないデータ削除が行われる場合があるため、[-y] オプションは非推奨としています。

8 すべてのファイルシステムパーティションをリストアするまで、手順 5 ～手順 7 を繰り返します。

9 すべてのパーティションのリストアが完了したら、アンマウントしディレクトリを削除します。

```
# ssh ＜IPアドレスまたはホスト名＞ umount /backup
# ssh ＜IPアドレスまたはホスト名＞ rm -rf /backup
```

10 以下のコマンドを実行し、システムを再起動します。

```
# reboot
```

11 **システムボリュームをリストアした場合は、続けて「8.3.8 ブートローダの設定（システムボリュームのリストア時）」（→ P.192）を実施してください。**

8.3.8 ブートローダの設定（システムボリュームのリストア時）

システムボリュームのリストア後に、併せてブートローダ（Grub）の設定が必要となります。

1 **レスキューモードでシステムを起動します。**

レスキューモードの起動方法については、「7.4 レスキューモードの起動（トラブル対応時）」（→ P.154）を参照してください。

2 **「Rescue」画面で既存のパーティションを自動的にマウントするかどうかのメッセージが表示されたら、[Continue] を選択します。**

リストアしたファイルシステムの "/"（ルート）パーティションが /mnt/sysimage にマウントできた場合には、結果が表示されるので、[OK] を選択します。

注 意

- **リストアした OS の /etc/fstab に記載しているデバイスのうち、未接続のデバイスがある場合、エラーが発生する可能性があります。エラーが発生した場合は、当該デバイスを記載している行をコメントアウトして、再度手順 1 から実行してください。**

3 **以下のコマンドを実行し、/mnt/sysimage や /mnt/sysimage/boot がマウントされていることを確認します。**

```
# df
```

4 **以下のコマンドを実行し、ルートパスをディスクに変更します。**

```
# chroot /mnt/sysimage
```

5 **以下のコマンドを実行し、ブートローダを Master Boot Record（以降、MBR と表記します）にインストールします。**

```
# /sbin/grub-install /dev/sda
```

/dev/sda は MBR が存在するデバイス名です。/dev/sda を指定してください。

6 **手順 2 で /etc/fstab を編集した場合は元に戻します。**

7 **以下のコマンドを実行し、chroot 環境を抜けます。**

```
# exit
```

8 **以下のコマンドを実行し、システムを再起動します。**

```
# reboot
```

参 考

- 手順 2 で /etc/fstab を編集した場合は、システムを停止後、必要なデバイスを接続してからシステムを起動してください。

用語解説

■ 修正ユリウス日

紀元前 4713 年 1 月 1 日正午（世界標準時による）からの日数であるユリウス通日（AJD）から 2400000.5 を引いたもの。現在使われているグレゴリオ暦から修正ユリウス日を以下の式で求めることができる。x を超えない最大の整数を [x] で表すと、グレゴリオ暦 y 年 m 月 d 日午前 0 時の修正ユリウス日は、以下のとおりである。

[365.25y]+[y/400]-[y/100]+[30.59(m-2)]+d-678912

ただし、1 月、2 月の場合は、前年の 13 月、14 月として計算する。

参考：2001 年 1 月 1 日の修正ユリウス日は 51910

■ トランキング

trunking

ネットワークのスイッチなどにおいて、隣接した複数のポートを多重化し、一本の高速リンクとして機能させること。

■ GDS

Global Disk Services

製品名は PRIMECLUSTER GDS。富士通が提供するシングルシステムでのストレージの冗長化を行うボリューム管理ソフトウェア。

■ GLS

Global Link Services

製品名は PRIMECLUSTER GLS。富士通が提供する、ネットワーク伝送路を冗長化し、通信の高信頼化を実現するソフトウェア。

■ GPS

Global Positioning System

地球を周回する 20 数個の衛星から発信される電波を利用し、緯度、経度、高度、時刻などを割り出すことができるシステム。セシウムなどを使った高精度な原子時計が装備されている。

■ GSPB

Giga-LAN SAS and PCI_Box connector Board

オンボード I/O、および PCI_Box 用 PCI Express インターフェースを搭載したボード。

■ master デバイス

bonding ドライバが構成した仮想的な NIC を示す。通常は bond ＜ n ＞という形式のインターフェース名。

■ MMB

ManageMent Board

筐体内の制御・監視、パーティション管理、システム初期化などを行うシステム制御ユニット。

■ MSTP

Multiple Spanning Tree Protocol

STP プロトコルを VLAN を意識して動作するように拡張したプロトコル。物理的には閉路を形成していても、VLAN 構成によっては VLAN 単位には閉路にならない場合などがあり、こういったケースでも VLAN 単位に閉路を検出して切断するためのプロトコルである。

■ NIC

Network Interface Card

コンピュータネットワーク内でコンピュータ間の通信を行うために使用されるハードウェアの 1 つ。

■ SAS

Serial Attached SCSI

SCSI 規格のうちの 1 つであり、コンピュータにハードディスクなどのデバイスを接続するためのインターフェースである。シリアル伝送を行う。

■ SASU

SAS ディスクユニット

内蔵ハードディスクドライブとそのコントローラーとしての PCI カードを搭載できるユニット。

■ slave デバイス

master デバイスを構成する個々の NIC を示す。通常の eth ＜ n ＞という形式のインターフェースが slave デバイスとして組み込まれる。

■ STP

Spanning Tree Protocol

スイッチでサポートされるプロトコルの 1 つ。冗長性のあるスイッチネットワークを構成すると、閉路が形成されることがあり、フレームがループすることがある。STP はスイッチ間でリンクの接続形態と状態を交換して閉路を形成するリンクを切断し、フレームのループを防止するプロトコルである。

■ UTC

Universal Time, Coordinated

全世界で公式な時刻を意味する。時刻には以下の種類がある。
・世界時（UT1）：天体観測を元に決められている時刻。
・国際原子時（TAI）：世界時（UT1）における 1958 年 1 月 1 日 0 時 0 分 0 秒からの経過時間を、高精度なセシウム原子時計でカウントして定めた時刻。
・協定世界時（UTC）：国際原子時（TAI）に世界時（UT1）とのずれを調整するための「うるう秒」を導入した時刻。現在は、グリニッジ標準時（GMT）に代わって標準になっている。

索　引

み

ゆ

ら

り

れ

ろ